

Université de Montréal

**Dynamic epigenetic changes in immune responses to  
infection in human dendritic cells**

par  
Alain Pacis

Programme de bio-informatique  
Faculté de médecine

Mémoire présenté à la Faculté des études supérieures et postdoctorales  
en vue de l'obtention du grade de M.Sc.  
en bio-informatique

Mai 2014  
© Alain Pacis, 2014

Université de Montréal  
Faculté des études supérieures

Ce mémoire intitulé :

**Dynamic epigenetic changes in immune responses to  
infection in human dendritic cells**

présenté par

**Alain Pacis**

a été évalué par un jury composé des personnes suivantes :

**Benoît Coulombe, Ph.D.**

président-rapporteur

**Luis B. Barreiro, Ph.D.**

directeur de recherche

**Noël Raynal, Ph.D.**

membre du jury

## Résumé

La méthylation de l'ADN est une marque épigénétique importante chez les mammifères. Malgré le fait que la méthylation de la cytosine en 5' (5mC) soit reconnue comme une modification épigénétique stable, il devient de plus en plus reconnu qu'elle soit un processus plus dynamique impliquant des voies de méthylation et de déméthylation actives. La dynamique de la méthylation de l'ADN est désormais bien caractérisée dans le développement et dans le fonctionnement cellulaire des mammifères. Très peu est cependant connu concernant les implications régulatrices dans les réponses immunitaires. Pour se faire, nous avons effectué des analyses du niveau de transcription des gènes ainsi que du profilage épigénétique de cellules dendritiques (DCs) humaines. Ceux-ci ont été faits avant et après infection par le pathogène *Mycobacterium tuberculosis* (MTB). Nos résultats fournissent le premier portrait génomique du remodelage épigénétique survenant dans les DCs en réponse à une infection bactérienne. Nous avons constaté que les changements dans la méthylation de l'ADN sont omniprésents, identifiant 3,926 régions différenciellement méthylées lors des infections par MTB (MTB-RDMs). Les MTB-RDMs montrent un chevauchement frappant avec les régions génomiques marquées par les histones associées avec des régions amplificatrices. De plus, nos analyses ont révélées que les MTB-RDMs sont activement liées par des facteurs de transcription associés à l'immunité avant même d'être infecté par MTB, suggérant ces domaines comme étant des éléments d'activation dans un état de dormance. Nos données suggèrent que les changements actifs dans la méthylation jouent un rôle essentiel pour contrôler la réponse cellulaire des DCs à l'infection bactérienne.

**Mots-clés:** Epigénétique, méthylation de l'ADN, modifications des histones, dynamique de la chromatine, régions amplificatrices, infection bactérienne, inflammation, bacille de Koch

# Abstract

DNA methylation is an important epigenetic mark in mammals. Although methylation at the 5' position of cytosine (5mC) is recognized as a stable epigenetic modification, it is becoming increasingly viewed as a more dynamic process that involves both active methylation and demethylation pathways. While the dynamics of DNA methylation has been well characterized in mammalian development and normal cellular function, little is known about its regulatory implications in immune responses. To that end, we performed comprehensive transcriptional and epigenetic profiling of primary dendritic cell (DC) samples from humans, before and after infection with *Mycobacterium tuberculosis* (MTB). Our results provide the first complete genomic portrait of the extensive epigenetic remodeling occurring in primary DCs in response to a bacterial infection. We found that active changes in DNA methylation are pervasive, identifying 3,926 MTB-induced differentially methylated regions (MTB-DMRs). MTB-DMRs show a striking overlap with genomic regions marked by histones associated with enhancer activity. ATAC-seq footprinting analysis revealed that regions that change methylation were actively bound by immune-related TFs prior to MTB-infection suggesting that these domains are likely to represent enhancer elements in a poised state. Our data suggests that active changes in DNA methylation play an essential and previously unappreciated role at controlling of the regulatory programs engaged by DCs in response to a bacterial infection.

**Keywords:** Epigenetics, DNA methylation, histone modifications, chromatin dynamics, enhancers, bacterial infection, inflammation, *Mycobacterium tuberculosis*

# Table of Contents

Résumé.....	iii
Abstract.....	iv
List of tables.....	vii
List of figures.....	viii
List of abbreviations .....	ix
Acknowledgments.....	xi
1 Introduction.....	1
1.2 DNA methylation.....	2
1.3 Patterns of global methylation in mammals.....	3
1.3.1 CpG islands.....	3
1.3.2 Gene body .....	4
1.3.3 Intergenic .....	5
1.3.4 Non-CpG methylation.....	5
1.4 Basic mechanism of DNA methylation .....	6
1.4.1 Writing DNA methylation .....	6
1.4.2 Erasing DNA methylation.....	8
1.4.3 Reading DNA methylation .....	10
1.5 Approaches to investigate DNA methylation .....	11
1.5.1 Bisulfite sequencing.....	11
1.5.2 Bisulfite microarrays.....	12
1.5.3 Enrichment-based methods.....	13
1.6 Crosstalk between DNA methylation and histone modifications .....	13
1.7 Dendritic cells in shaping the immune response.....	15
1.8 Research objectives.....	18
2 Experimental Approach .....	19
2.1 General molecular biology.....	19
2.2 Bioinformatics analyses .....	22
3 Results.....	27
3.1 SUMMARY .....	29

3.2 INTRODUCTION .....	30
3.3 RESULTS .....	32
3.4 DISCUSSION .....	37
4 Conclusion .....	51
Bibliography .....	53

## List of tables

*Dynamic epigenetic changes in immune responses to infection in human dendritic cells*  
*[Manuscript to be submitted]*

Table S1 Correlation of Methylation Levels Between DC Samples .....	48
---	----

## List of figures

Figure 1.1 Schematic of epigenetic markers involved in gene regulation .....	2
Figure 1.2 Establishment of DNA methylation patterns in mammals .....	7
Figure 1.3 A complete pathway for dynamic cytosine modifications .....	9
Figure 1.4 Readers of methylation signal and their potential mechanism in gene repression ..	11
Figure 1.5 Bisulfite conversion .....	13
Figure 1.6 Transcriptional regulation is achieved through crosstalk between DNA methylation and histone modifications .....	15
Figure 1.7 Dendritic cell maturation upon antigen encounter .....	17
<i>Dynamic epigenetic changes in immune responses to infection in human dendritic cells</i> <i>[Manuscript to be submitted]</i>	
Figure 1. DC Methylome is Dynamically Remodeled during MTB Infection .....	40
Figure 2. MTB-DMRs Have Chromatin Characteristics of Enhancers .....	41
Figure 3. MTB-DMRs are Enriched for Immune-Specific Transcription Factor Footprints ..	43
Figure 4. Differential Methylation is Coupled to Differential Gene Expression.....	44
Figure S1. General Workflow of Study .....	45
Figure S2. Global Trends of DNA Methylation in DCs .....	47
Figure S3. Genomic Features of MTB-DMRs.....	48
Figure S4. Chromatin Dynamics at Hypermethylated Regions .....	49
Figure S5. Transcriptional Profiles of DC Samples.....	50



## List of abbreviations

5caC	5-carboxyl cytosine
5fC	5-formyl cytosine
5hmC	5-hydroxymethyl cytosine
5mC	5-methyl cytosine
AM	Active modification
AR	Active restoration
ATAC-seq	Assay for transposase-accessible chromatin using sequencing
BER	Base excision repair
bp	Base pair
ChIP-seq	Chromatin immunoprecipitation
cDNA	Complimentary DNA
CGI	CpG island
CpG	Cytosine-Guanine dinucleotide
DC	Dendritic cell
DMR	Differentially methylated region
DNA	Deoxyribonucleic acid
TDG	Thymine DNA glycosylase
Dnmt	DNA methyltransferase
eRNA	Enhancer RNA
GO	Gene Ontology
H3K4me1	Histone 3 lysine 4 monomethylation
H3K4me3	Histone 3 lysine 4 trimethylation
H3K9me3	Histone 3 lysine 9 trimethylation
H3K27ac	Histone 3 lysine 27 acetylation
H3K27me3	Histone 3 Lysine 27 trimethylation
H3K36me3	Histone 3 Lysine 36 trimethylation
HDAC	Histone deacetylase
kb	Kilobase
mRNA	Messenger RNA

MethylC-seq	Whole genome shotgun bisulfite sequencing
MTB	<i>Mycobacterium tuberculosis</i>
PD	passive dilution
PCR	Polymerase chain reaction
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
TB	Tuberculosis
Tet	Tet methylcytosine dioxygenase
TF	Transcription factor
TSS	Transcription start site

# Acknowledgments

I would like to thank my supervisor, Luis Barreiro for giving me the opportunity to work on an exciting project and for his invaluable assistance, support and guidance throughout the thesis.

I also wish to acknowledge my collaborators for their excellent experimental contribution and scientific interest.

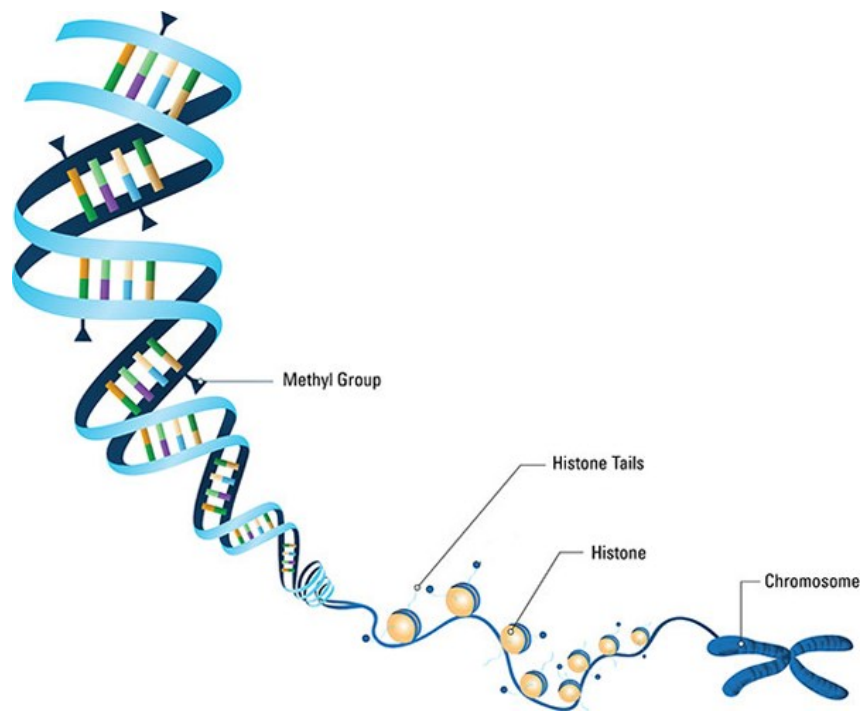
I am grateful to all past and present members of the Barreiro lab for their help in different parts of the work and for creating a scientifically motivating environment. I am fortunate to be a part of an amazing group who made this period as enjoyable as possible.



# 1 Introduction

## 1.1 Epigenetics

Despite the fact that every cell in a given multicellular organism contains the same genetic information, each cell exhibits different functions and morphologies. Knowing the nucleotide sequence alone is only a small part of the puzzle and that the answer lies in the epigenetic regulation of genes. To understand epigenetics requires an understanding of chromatin structure. DNA is packaged into a highly organized and repeated protein-DNA complex called nucleosomes, which form chromatin. In addition, our genome is embellished with many covalent modifications, which together make up our “epigenome”. These alterations are called epigenetic because they do not involve mutations in the underlying DNA sequence. DNA methylation and post-translational modifications of core histones are the best-known forms of epigenetic process (see **Figure 1.1**). Combination of these modifications characterizes the chromatin configuration and the accessibility of the DNA to the transcription machinery and thus, governs the transcriptional regulation of the expression of genes (Fischle, 2008; Cedar and Bergman, 2009). In the past few decades, significant progress has been made toward our understanding of the importance of epigenetic mechanisms in a wide range of biological processes such as organism development and disease progression.



**Figure 1.1 Schematic of epigenetic markers involved in gene regulation.** The two most studied components of the “epigenetic code” are DNA methylation and histone tail modification. <http://www.thermoscientificbio.com/applications/epigenetics/> (2014)

## 1.2 DNA methylation

DNA methylation is the covalent addition of a methyl group to the fifth carbon of the cytosine base that leads to the formation of 5-methylcytosine (5mC). 5mC has been known to be associated with a transcriptionally repressed chromatin state, particularly in long-term gene silencing. In 1979, McGhee and Ginder compared the methylation status of a single beta-globin locus between cells that did and did not express this gene (McGhee and Ginder, 1979). At that time, they used restriction enzymes to distinguish between methylated and unmethylated DNA. They showed that the beta-globin locus is essentially unmethylated in cells that expressed beta-globin. In addition, adding different chemicals that destroyed methyl groups such as, 5-azacytidine, often turned such genes back on. Shortly after this pioneering work, 5-Aza-2'-deoxycytidine has been shown to severely inhibit the action DNA methyltransferases to normally methylate DNA cytosine residues (Jones *et al.*, 1983). Using this chemical product, scientists compared cells before and after treatment to see how the loss of methylation impacts

gene expression by looking at changes in cellular phenotypes. These experiments inspired investigators to better understand how exactly DNA methylation impacts gene expression. For the past years, DNA methylation has been extensively studied for its role in various biological processes including genomic imprinting, inactivation of one X chromosome in females, suppression of mobile genetic elements, stem cell differentiation and embryonic development. Moreover, aberrant methylation leads to cancerous growth and several other diseases.

### **1.3 Patterns of global methylation in mammals**

Patterns of DNA methylation vary among different organisms. In mammals, DNA methylation occurs predominantly at CpG (cytosine-phosphate-guanosine) dinucleotides but is also found at non-CpG sites at a much lesser extent (Lister *et al.*, 2009; Laurent *et al.*, 2010; Hon *et al.*, 2013; Ziller *et al.*, 2013). Overall, mammalian genomes are very rich in DNA methylation with the exception of regions called CpG islands (described below).

#### **1.3.1 CpG islands**

The most prominent feature of DNA methylation is found within CpG islands (CGIs) near transcription start sites (TSS) of genes. CGIs are stretches of DNA that have a higher CpG density than the rest of the genome but are commonly devoid of methylation (Bird *et al.*, 1985). Computational analysis estimated 28,890 CGIs in our genome (Lander *et al.*, 2001), which are highly conserved between mice and humans (Deaton and Bird, 2011). Interestingly, the majority of gene promoters, particularly of housekeeping genes, reside within CGIs (Gardiner-Garden and Frommer, 1987; Saxonov *et al.*, 2006; Illingworth *et al.*, 2010). The location and preservation of CGIs throughout evolution strongly suggests that these regions possess a functional importance.

Significant evidence indicates that CGIs have been evolutionarily conserved to promote gene expression by modulating the chromatin structure and transcription factor (TF) binding. One of the common features of CGIs is that they contain more permissive chromatin than other stretches of DNA and therefore promote gene expression (Tazi and Bird, 1990; Ramirez-Carrozzi *et al.*, 2009; Choi, 2010). Despite their lack of common promoter elements (i.e., TATA boxes), CGIs enhance the accessibility of DNA and promote TF binding.

Furthermore, there have been many findings of strong inverse correlations between promoter activity and DNA methylation, suggesting that promoter methylation exert a regulatory function on gene activity. This is best exemplified in the process of genomic imprinting where methylation at CpG islands results in stable silencing of gene expression (Li *et al.*, 1993; Kaneda *et al.*, 2004). In cancer phenotypes, abnormal gains of methylation in normally unmethylated promoter CpG islands of tumor suppressor genes lead to stabilization of transcriptional repression and loss of gene function (Jones and Baylin, 2002; Robertson, 2005). CGIs also undergo differential methylation and regulate gene expression during development and differentiation (Weber *et al.*, 2007; Fouse *et al.*, 2008; Meissner *et al.*, 2008).

### **1.3.2 Gene body**

As mammalian genomes are globally methylated at CpG sites, it is not surprising that DNA methylation is also found across the gene. The ubiquity of DNA methylation makes it difficult to assess whether DNA methylation is targeted in a specific manner or if it's the default state of CpG dinucleotides. Cytosine methylation in the promoter region of genes represses transcription and yet the bodies of active genes are extensively methylated. Indeed, there has been evidence suggesting that DNA methylation of the gene body is associated with a higher level of gene expression in dividing cells (Ball *et al.*, 2009; Aran *et al.*, 2011). Also, in X chromosome inactivation studies, Hellman and Chess compared the DNA methylation levels on the active ( $X_a$ ) and inactive ( $X_i$ ) X chromosomes and found reduced methylation specifically over gene bodies on  $X_i$  (Hellman and Chess, 2007). On the other hand, in postmitotic cells such as the brain, gene body methylation is not associated with increased gene expression (Guo *et al.*, 2011). As there is increasing evidence that methylation of DNA in promoters suppresses gene expression, the role of gene body methylation still remains unclear.

Recently, Maunakea *et al.* showed that a number of methylated CpG islands are located in intragenic and intergenic regions using RNA markers of transcription initiation and histone markers (Maunakea *et al.*, 2010). By further separating CGIs into intragenic and intergenic categories, they found that intragenic CGIs are especially prone to methylation, supporting the role of intragenic methylation in regulating alternative promoter activity. Another hypothesis is that gene body methylation enhances accurate splicing of pre-mRNA supported by the



experimental data showing a higher level of methylation in exons compared to introns and that spliced exons tend to display lower levels of methylation (Maunakea *et al.*, 2013). Gene body methylation also potentially plays a role in silencing repetitive DNA elements (i.e., LINE, Alu, etc.) (Yoder *et al.*, 1997). CpG sites in these repetitive regions are largely methylated and are believed to be suppressing their transposition activity.

### **1.3.3 Intergenic**

The potential association of DNA methylation beyond promoters, more specifically at distal regulatory regions, has been overlooked because of many reported abnormal changes in CpG island methylation situated in cancer patients. Distal regulatory elements (enhancers, silencers and insulators) are short regions of DNA of low CpG density located at variable distances from the TSS of genes. These regions are found almost exclusively within regions of accessible chromatin and play a key role in increasing or decreasing transcription of target genes. Recent genome-wide analyses document strong anti-correlation between transcription factor and coactivator binding and DNA methylation density (Lister *et al.*, 2009; Stadler *et al.*, 2011; Neph *et al.*, 2012) at enhancer. Moreover, these regions are neither methylated nor unmethylated but exhibit a more variable methylation illustrating that DNA methylation at regulatory elements is much more dynamic than previously appreciated.

### **1.3.4 Non-CpG methylation**

As previously mentioned, the majority of cytosines in non-CpG context are devoid of methylation in mammals. Although this preference of methylation towards the CpG context is unknown, there is evidence that non-CpG DNA methylation is relatively abundant in oocytes, pluripotent embryonic stem cells, and mature neurons (Lister *et al.*, 2009; Xie *et al.*, 2012; Lister *et al.*, 2013; Shirane *et al.*, 2013; Guo *et al.*, 2014). However, the function of non-CpG methylation remains unclear. Non-CpG methylation is likely to arise from the same family of *de novo* methyltransferases since there are no recorded homologs of Dnmts to methylate non-CpG cytosines.

## 1.4 Basic mechanism of DNA methylation

It is clear that understanding how genome-wide methylation patterns are established requires elucidating the mechanisms involved in regulating DNA methylation. Thanks to decades of research, many of the proteins and mechanisms involved in DNA methylation have already been identified. Processes that regulate DNA methylation are essentially broken down into three classes: writers, erasers, and readers. Writers are the enzymes that catalyze the addition of methyl groups onto cytosine residues. Erasers modify and remove methyl groups. Finally, readers recognize and bind to methyl groups to ultimately mediate changes in gene expression.

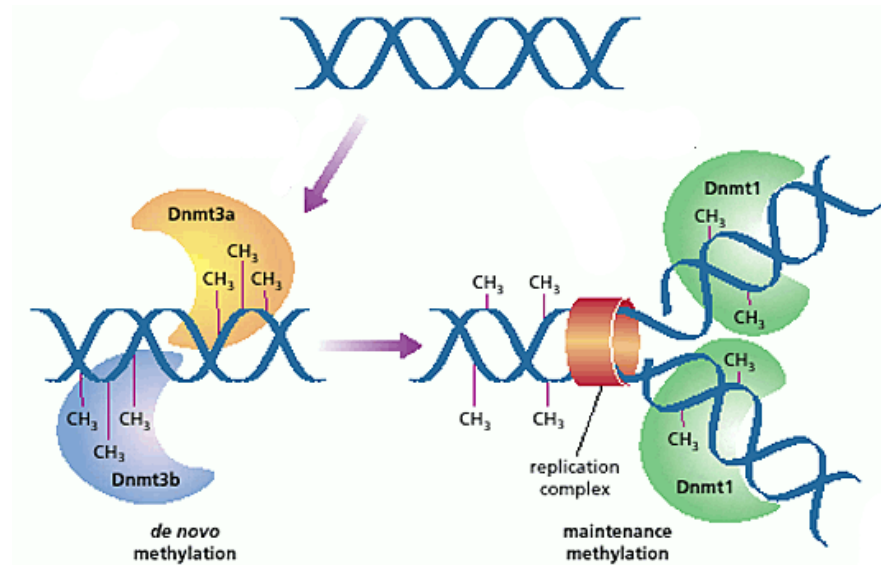
### 1.4.1 Writing DNA methylation

DNA methylation occurs by the addition of a methyl group from S-adenosylmethionine to cytosine with the help of DNA methyltransferases (Dnmts) (Schermelleh *et al.*, 2005). As shown in **Figure 1.2**, there are two types of DNA methyltransferase activities in mammals: *de novo* and maintenance methylation, which are achieved by Dnmt1 and Dnmt3 respectively. Dnmt1 was the first eukaryotic DNA methyltransferase to be discovered (Bestor *et al.*, 1988). Dnmt1 seemed to be responsible only for maintaining methylation after each round DNA replication, which led to the assignment of Dnmt1 as a maintenance DNA methyltransferase (Bestor, 2000; Goyal *et al.*, 2006). This is supported by findings showing that Dnmt1 co-localizes with the replication machinery (Leonhardt *et al.*, 1992; Schermelleh *et al.*, 2007). At replication sites, hemimethylated DNA is formed when the newly synthesized unmethylated strand pairs with the methylated template strand. Strikingly, while virtually all of methylated CpG sites are methylated in both strands, 98% of methylated CHG sites are highly asymmetrical with only one of strands being methylated (Lister *et al.*, 2009). This suggests that Dnmt1 recognizes its substrate cytosine residue only if a guanine residue is beside it.

In addition to Dnmt1, the mammalian genome encodes two functional Dnmt3 methyltransferases, namely Dnmt3a and Dnmt3b. In the late 1990's, Okano *et al.* have reported both enzymes as *de novo* Dnmts since they were found to methylate CpG dinucleotides in vitro without preference for hemimethylated DNA (Okano *et al.*, 1999). Although Dnmt3a and Dnmt3b show considerable functional redundancy in early developmental stages, they have different expression profiles in distinct cell types. Moreover, Dnmt3b appears to be specialized

in particular parts of the chromosome as it engages methylation only at the centromeric region (Xu *et al.*, 1999). It has also been proposed that there is a possible cooperation between the *de novo* and the maintenance Dnmts (Siedlecki and Zielenkiewicz, 2006). Dnmt3a and Dnmt3b may also participate in the maintenance of methylation by restoring methylation at cytosine residues which have been overlooked by Dnmt1 during replication.

There is a third homolog in the Dnmt3 family found only in germ cells, called Dnmt3L (DNA methyltransferase 3-like). Although this protein has been shown to not possess methyltransferase activity, it is essential as a regulatory cofactor of Dnmt3a and Dnmt3b (Goll and Bestor, 2005). The family Dnmt2 cytosine methyltransferases is the most strongly conserved among all known cytosine methyltransferase and it is ubiquitously expressed in most human and mouse tissues. What makes Dnmt2 of enigmatic nature is that it lacks methyltransferase activity and does not seem to act as a regulatory factor like Dnmt3L – mice deficient of Dnmt2 does not display any overt phenotype (Siedlecki and Zielenkiewicz, 2006).



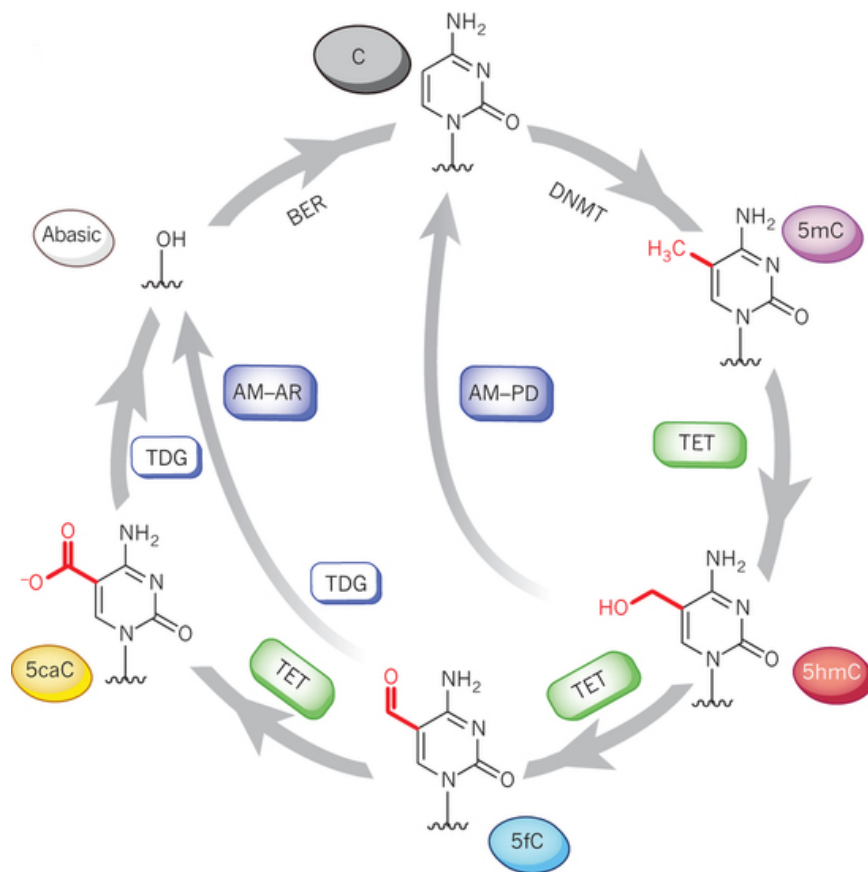
**Figure 1.2 Establishment of DNA methylation patterns in mammals.** Dnmt3a and Dnmt3b function as *de novo* methyltransferases whereas Dnmt1 is targeted to replication forks to introduce methyl groups to cytosines on the newly synthesized DNA strand corresponding to the methylated cytosines in the template strand. *Image adapted from (Reik et al., 1999).*

### 1.4.2 Erasing DNA methylation

DNA methylation is relatively stable compared with other epigenetic marks such as histone modifications. Nevertheless, loss of DNA methylation, or DNA demethylation, has been observed in different biological contexts. DNA demethylation is the process of removal of a methyl group from nucleotides in DNA and it may take place in a passive or active fashion. Passive DNA demethylation takes place in dividing cells. As Dnmt1 maintains DNA methylation during cell replication, its absence allows newly synthesized DNA strands to be devoid of methylation. Active DNA demethylation occurs via direct removal of a methyl group independently of DNA replication and therefore can take place in both dividing and non-dividing cells. So far, there is no known mechanism in mammalian cells that can cleave the strong covalent carbon-to-carbon bond that connects cytosine to a methyl group. Instead, active demethylation occurs through a series of chemical reactions that revert 5mC back to C. A series of recent discoveries has brought clarity on our understanding of active DNA demethylation.

Until recently, the only known covalent epigenetic modification on DNA was methylation at position 5' of cytosine. A landmark discovery by Tahiliani *et al.* was made showing that 5mC is oxidized to 5-hydroxymethylcytosine (5hmC) by the enzyme ten-eleven translocation (Tet) family proteins (Tahiliani *et al.*, 2009). More importantly, work from the same group have shown that Tet proteins and 5hmC may be involved in DNA demethylation – overexpression of Tet1 leads to a decrease in 5mC levels. Since the discovery of Tet, 5hmC has taken on a new central role in DNA demethylation. Given our current understanding, active demethylation involving Tet fall into two groups, which both initially involve active modification (AM) of 5mC to generate 5hmC (see **Figure 1.3**) (Kohli and Zhang, 2013). In the process of passive dilution (PD), unmodified C is regenerated through DNA replication since Dnmt1 does not recognize 5hmC and therefore cannot maintain it (Inoue *et al.*, 2011; Inoue and Zhang, 2011). Alternatively, active restoration (AR) requires further enzymatic modification of 5hmC to regenerate unmodified C. Specifically, Tet can further oxidize 5hmC, yielding 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (He *et al.*, 2011; Ito *et al.*, 2011). 5fC or 5caC is subsequently excised by thymine DNA glycosylase (TDG) generating an abasic site as part of the base excision repair (BER) process (Fromme and Verdine, 2004). Ideally, AM-AR has the advantage of achieving rapid conversion of 5mC to unmodified C and

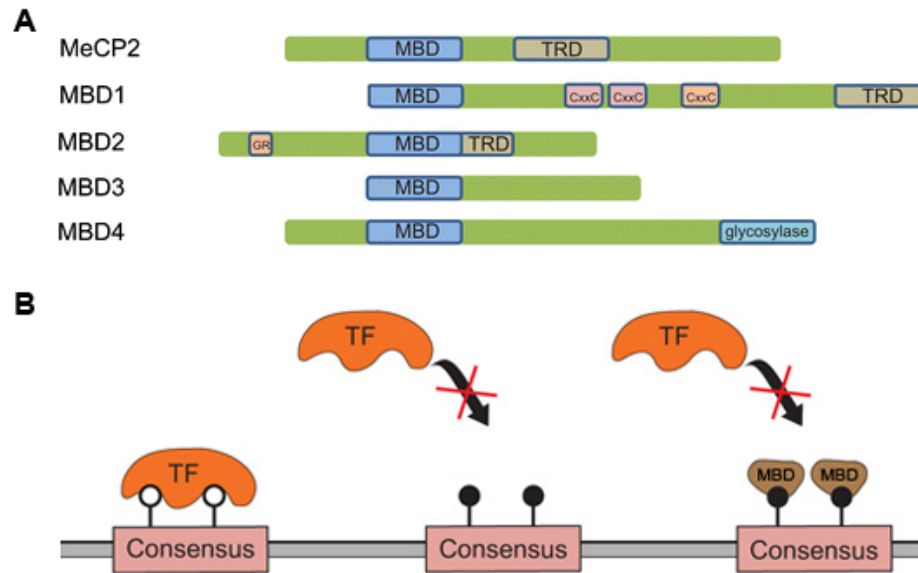
therefore seems particularly well suited to locus-specific demethylation processes that require a rapid response to environmental stimuli.



**Figure 1.3 A complete pathway for dynamic cytosine modifications.** 5mC bases are introduced by DNA methyltransferase (DNMT) enzymes. 5mC is further oxidized to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) during active DNA demethylation by Tet family proteins. In the pathway of active modification followed by passive dilution (AM-PD), 5hmC is diluted in a replication-dependent manner to regenerate unmodified C. In the pathway of AM followed by active restoration (AM-AR), 5hmC is further oxidized by TET proteins to 5fC and then 5caC, which is then excised by TDG and repaired by the BER pathway into an unmodified cytosine, generating an abasic site. *Image adapted from (Kohli and Zhang, 2013).*

### 1.4.3 Reading DNA methylation

There are potentially multiple ways in which DNA methylation can decrease transcription levels or completely turn off genes. One simple way this is accomplished is by directly preventing the binding of the transcriptional machinery or transcription factors to their respective regions by the methyl groups themselves (Hark *et al.*, 2000; Jaenisch and Bird, 2003). Alternatively, repression can be achieved via proteins that specifically recognize and bind 5mC. The identification of proteins that can read the methylation signal has shed light on how DNA methylation play a repressive role in gene expression. The first characterized methylcytosine-binding protein family is the methyl-CpG binding domain (MBD) family which consists of five members, namely MBD1, MBD2, MBD3, MBD4 and MeCP2. These proteins each contain of a conserved MBD domain that confers a high affinity for methylated CpG sites (Nan *et al.*, 1993). As illustrated in **Figure 1.4**, three of these proteins (MeCP2, MBD1 and MBD2) also contain a transcriptional repression domain (TRD) that allows these proteins to interact with a variety of repressor complexes and participate in methylation-dependent repression of transcription (Meehan *et al.*, 1989; Nan *et al.*, 1993). As it will be discussed in more detail later, TRDs are known to associate with various histone modification enzymes and chromatin remodeling complexes. The second family of methyl-binding proteins binds to methylated DNA via a zinc-finger domain and is composed of Kaiso, ZBTB4, and ZBTB38 (Prokhortchouk *et al.*, 2001; Filion *et al.*, 2006). Finally, there is the family of UHRF (ubiquitin-like, containing PHD and RING finger domain) proteins that includes, UHRF1 and UHRF2. These are multidomain proteins that bind methylated cytosines via a SET- and RING-associated DNA-binding domain (Hashimoto *et al.*, 2008; Hashimoto *et al.*, 2009).



**Figure 1.4 Readers of methylation signal and their potential mechanism in gene repression.** **A)** Graphical representation of the domains in the proteins that recognize and bind specifically methylcytosines (Bogdanovic and Veenstra, 2009). **B)** Transcription factor (TF) is able to bind to unmethylated cytosine (white circles) whereas methylated cytosines (black circles) or methylcytosine-binding domain (MBD) proteins interfere with TF binding. *Image adapted from* (Trzyna *et al.*, 2012).

## 1.5 Approaches to investigate DNA methylation

Over the past decade, epigenetics has grown dramatically and rapidly has become an exciting theme in genetics and bioinformatics. New methods have emerged to detect and measure DNA methylation, making it possible to map methylation patterns genome-wide, at a high-resolution and in a large number of samples (Harris *et al.*, 2010; Laird, 2010; Bock, 2012). There are three methods commonly used to assess DNA methylation: 1) bisulfite sequencing, 2) bisulfite microarrays, and 3) enrichment-based methods.

### 1.5.1 Bisulfite sequencing

Methylated cytosine has roughly the same base-pairing characteristics as unmethylated cytosine, and is thus indistinguishable by standard sequencing approaches. To overcome this,

genomic DNA can be treated with sodium bisulfite (Clark *et al.*, 1994; Clark *et al.*, 2006) (Clark *et al.*, 1994). Under appropriate conditions, sodium bisulfite causes the specific deamination of cytosine through a sulfonated intermediate, and its conversion to uracil. While unmethylated cytosine residues are converted to uracil, methylated cytosine residues are protected from this conversion and therefore remains intact. PCR amplification of converted DNA replaces the uracil with thymine. A potential issue with bisulfite analysis is that it depends on the complete conversion of unmethylated cytosines. It is essential to ensure that bisulfite-treated samples have been completely converted before utilizing them in high-throughput applications. Key advantages of this technology are its comprehensive genomic coverage, high quantitative accuracy and excellent reproducibility. Disadvantages include the high cost of whole-genome re-sequencing and the difficulty of discriminating between 5mC and 5hmC.

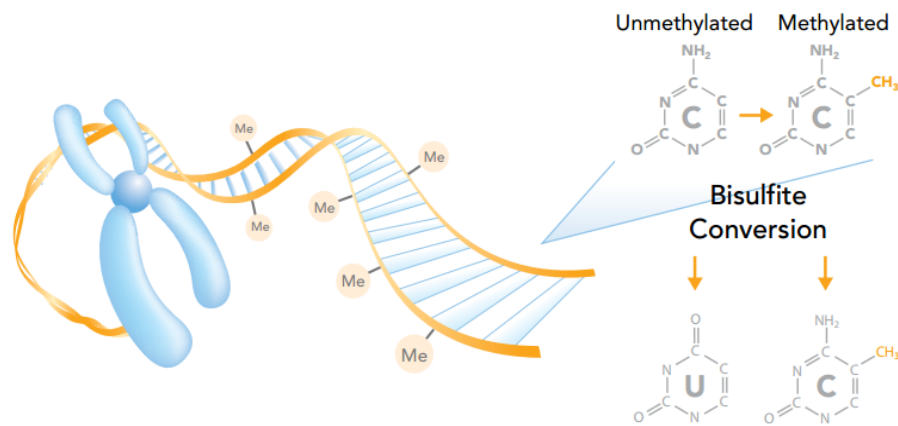
### **1.5.2 Bisulfite microarrays**

Bisulfite treatment in combination with specially designed genotyping microarrays makes it possible to measure DNA methylation levels at a preselected fraction of Cs throughout the genome. Bisulfite-converted DNA is assayed with two primers, each labeled with a different fluorescent dye. One primer is designed to hybridize if the cytosine is methylated (and unconverted), whereas the other will only hybridize to a converted sequence. The two primers are used in a PCR reaction with a locus-specific methylation-insensitive primer. This approach provides less coverage than other array-based methods, and necessitates the development and evaluation of a large set of selective primers, thus limiting its utility for *de novo* genome analysis. Bisulfite microarrays are commercially available only for the human genome. It also cannot distinguish 5mC from 5hmC. The strength of the technique is that it provides quantitative evaluation of specific cytosines and can process many samples in parallel. Therefore, this method is well suited to compare a set of known methylated loci among a large number of cell lines or individuals to ascertain methylation polymorphisms.



### 1.5.3 Enrichment-based methods

Methylated DNA immunoprecipitation (MeDIP), involves isolating methylated DNA using an antibody directed against 5-methylcytosine (Sorensen and Collas, 2009). After immunoprecipitation, the methylated and unmethylated part of the sample can be differentially labeled and applied to microarray (MeDIP-chip) or sequenced (MeDIP-seq). The two key advantages of enrichment-based methods are the relatively low cost of achieving genome-wide coverage (albeit with a low statistical power in CpG-poor genomic regions) and the ability to distinguish between different forms of DNA methylation: for example, using antibodies that specifically recognize 5hmC but not 5mC. However, these advantages come at the cost of relatively low resolution and high susceptibility to experimental biases.



**Figure 1.5 Bisulfite conversion.** The methyl group covalently attached to the 5' position of cytosine protects against bisulfite conversion. Unmethylated cytosines are converted to uracil, but not methylated cytosine.

[http://www.illumina.com/Documents/products/datasheets/datasheet\\_dna\\_methylation\\_analysis.pdf](http://www.illumina.com/Documents/products/datasheets/datasheet_dna_methylation_analysis.pdf) (2014)

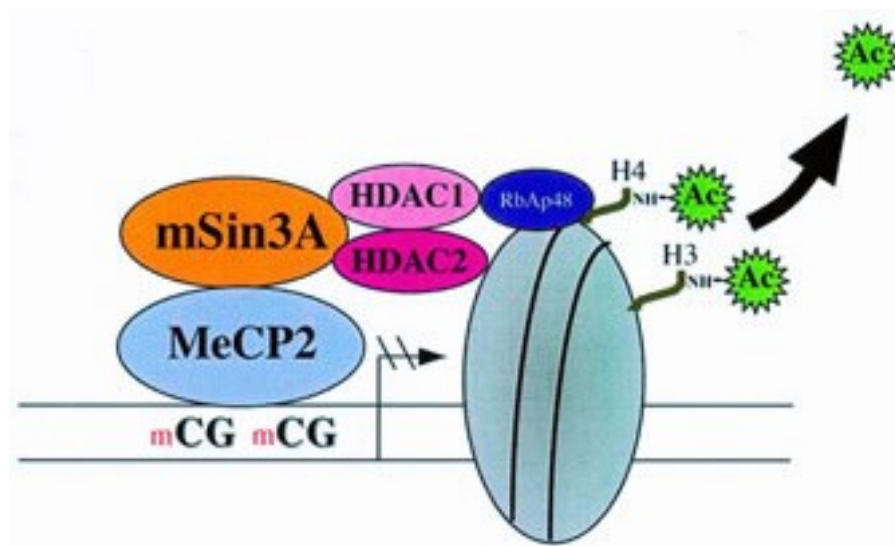
### 1.6 Crosstalk between DNA methylation and histone modifications

Transcription is ultimately regulated by the interaction of multiple epigenetic mechanisms, particularly DNA methylation and histone modifications, which cooperate to activate or silence gene expression. There is a wide range of different modifications in the carboxy- and

amino-terminal tails of histone proteins, including acetylation, methylation, phosphorylation, ubiquitylation, and SUMOylation (Berger, 2007).

Past studies have shown that there is an intrinsic link between DNA methylation and the methylated state of histone protein tails: Dnmt3L interacts with unmethylated H3K4 through its N-terminus and with Dnmt3a through its C terminus (Jia *et al.*, 2007; Ooi *et al.*, 2007). According to this model, patterns of methylation of H3K4 (including mono-, di- or trimethylation – H3K4me) are already established before *de novo* methylation of the embryonic genome by Dnmt3 family. If Dnmt3L recognizes that the histone moiety is methylated, Dnmt3 cannot bind and the underlying DNA region is therefore protected from *de novo* methylation. It is important to note however that the relationship between DNA methylation and histone methylation can work in both directions: DNA methylation itself can serve as a template for histone methylation. During cell division, the conserved methylation pattern in the nascent DNA strand helps reconstruct the patterns of H3K4me3 after replication. This is supported by the loss of H3K4me3 when CXXC finger protein 1 (Cfp1) protein, a CpG-binding protein known to be a component of histone methyltransferase complex, is depleted (Thomson *et al.*, 2010).

MeCP2 has been shown to possess a transcriptional repressor domain that binds the corepressor mSin3A (Razin, 1998). This corepressor protein constitutes the core of a multiprotein complex that includes histone deacetylases (HDAC1 and HDAC2) which in turn, removes acetyl groups from lysine and arginine amino acids located in the histone tails (see **Figure 1.6**). Consequently, an increase in the positive charge occurs due to the amine groups of lysine and arginine residues in the histone tails. The positively charged histone and the negatively charged DNA backbone (due to phosphate ions) encourage high-affinity binding leading to condensing of the DNA around the histones and hindering the access of transcription factors to the promoter. Overall, this shows the close link between DNA methylation and histone modifications in regulating gene expression.



**Figure 1.6 Transcriptional regulation is achieved through crosstalk between DNA methylation and histone modifications.** MeCP2 recognizes and binds to methylated CpG residues (mCG) and recruits corepressor mSin3A which is a multiprotein complex that includes histone deacetylases (HDAC1 and HDAC2). HDACs removes acetyl groups from H3 or H4 histone tails. *Image adapted from (Razin, 1998).*

### 1.7 Dendritic cells in shaping the immune response

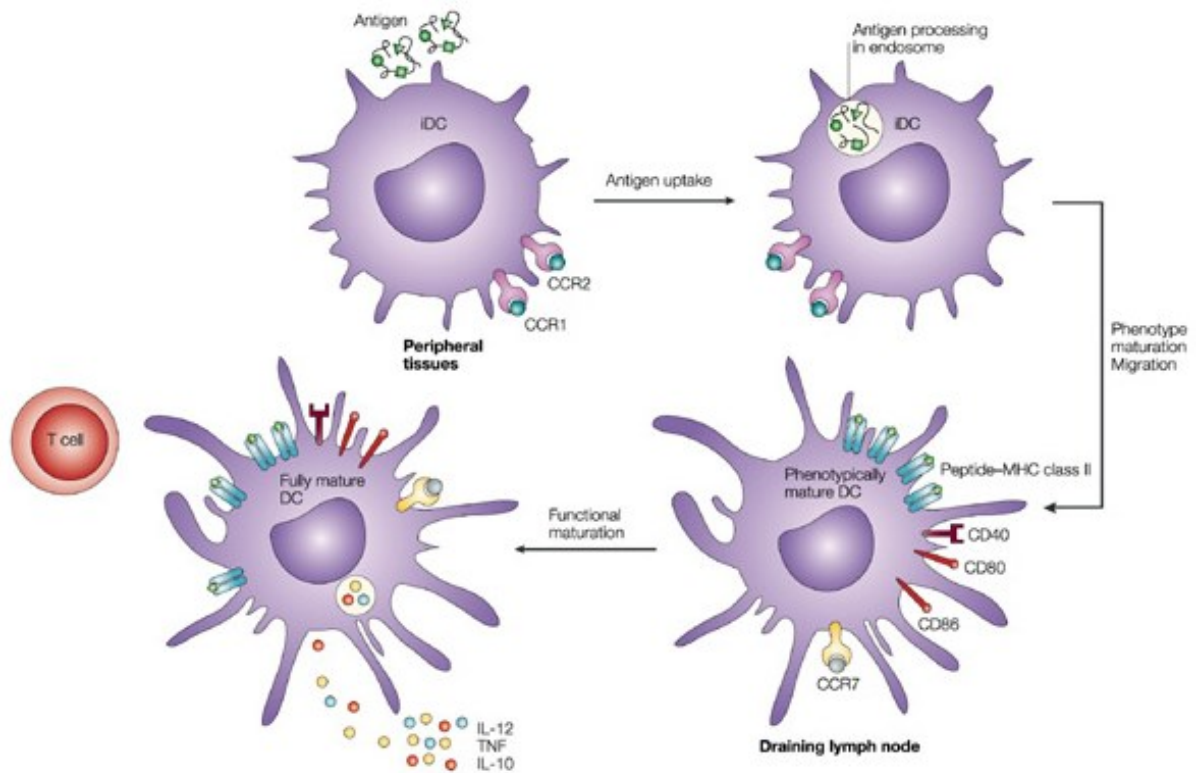
Dendritic cells (DCs) are professional antigen-presenting cells that have a central role in T cell activation and in initiation of adaptive immune responses. DCs express a number of pattern recognition receptors (PRRs) such as Toll-like receptors (TLRs) which recognize a wide array of pathogen-associated molecular patterns (PAMPs) and discriminate between self and non-self molecules (Kapsenberg, 2003). As immature cells specialized in antigen uptake and processing, DCs reside in and traffic through non-lymphoid peripheral tissues, continuously surveying the environment for invading microorganisms. As immature DCs capture antigens by endocytosis/phagocytosis, they undergo major changes in gene expression programs, evolving from immature, antigen-capturing cells to mature, antigen-presenting, T cell-priming cells. This process of DC maturation, in general, involves down-regulation of antigen internalization, a redistribution of major histocompatibility complex (MHC) molecules from

intracellular endocytic compartments to the DC surface, an increase in the surface expression of costimulatory molecules, secretion of chemokines and cytokines, and surface expression of adhesion molecules and chemokine receptors (Tan and O'Neill, 2005) (see **Figure 1.7**).

To present foreign antigens to naïve T cells, DCs must migrate from inflamed or injured peripheral tissues to the closest draining lymph nodes through afferent lymphatic vessels. Migration of maturing DCs from the periphery into lymphoid tissues are coordinated by chemokines that interact with corresponding receptors on DCs (Alvarez *et al.*, 2008). For example, immature DCs express CC-chemokine receptor 1 (CCR1), CCR2, CCR5 and CXC-chemokine receptor 1 (CXCR1) and are attracted to non-lymphoid tissues by their respective ligands, which are expressed constitutively or at inflammatory sites. DC maturation results in the downregulation of expression of these chemokine receptors and the upregulation of CCR7 expression. Expression of CCR7 switches DC responsiveness to its ligands, CC-chemokine ligand 19 (CCL19) and CCL21, that guide migration to secondary lymphoid organs. Maturation of DCs also induces the production of CCL22, CCL17 (i.e. chemokines that attract CCR4-expressing T cells), and CCL18. DC production of the chemokine CXCL16, in T cell-rich areas of lymphoid organs, may also function in promoting interaction between DCs and cytotoxic T cells.

DCs are capable of processing antigens and present peptide in the context of either MHC class I or II molecules, which interact with and stimulate cytotoxic T lymphocytes and T helper cells, respectively. As DCs mature, they acquire the properties necessary to form and transport peptide-loaded MHC complexes to the cell surface. Antigen transport to the cell surface coincides with increased expression of costimulatory molecules, such as CD40, CD80 and CD86 (Tan and O'Neill, 2005). These molecules amplify T cell receptor signaling and promote T cell activation. Moreover, the soluble cytokine profile secreted by DCs varies with the different stages of DC development and maturation thus influencing the different effector functions characteristic of immature vs. mature DCs (de Saint-Vis *et al.*, 1998). A wide variety of cytokines may be expressed by mature DCs and the exact cytokine repertoire expressed will depend on the nature of the stimulus, maturation stage of the DC and the existing cytokine microenvironment. The distinct cytokine patterns released by mature DCs ultimately

determine their Th1/Th2 polarizing capacities. For example, antigens that prime DCs to secrete IL-12 will typically induce Th1 differentiation.



**Figure 1.7 Dendritic cell maturation upon antigen encounter.** Maturation of dendritic cells, in response to antigen, leads to the redistribution of MHC class II molecules and MHC–peptide complexes from within the endocytic system to the cell surface, the production of several cytokines and membrane associated T cell stimulatory molecules, and the remodeling of expressed chemokine receptors. These changes allow dendritic cells to migrate to draining lymph nodes and induce antigen-specific immune response by activating T cells. *Image adapted from* (Hackstein and Thomson, 2004).

## **1.8 Research objectives**

Immune responses to infection are characterized by pervasive changes in gene expression with host cells mobilizing hundreds of genes involved in immune-related processes. While the dynamics of DNA methylation has been well-characterized in mammalian development and normal cellular function, little is known about its regulatory implications in the immune response. We hypothesized that widespread changes in DNA methylation may influence the expression of genes specifically involved in immune responses to *Mycobacterium tuberculosis* (MTB) infection. Our first aim is to identify genome-wide changes in DNA methylation between MTB-infected and non-infected dendritic cells (DCs). The second aim is to characterize the putative regulatory potential of changes in methylation. The third aim is to examine the interplay between DNA methylation and different histone marks. Finally, we aim to verify if dynamic DNA methylation is linked to MTB-induced changes in gene expression.

## 2 Experimental Approach

### 2.1 General molecular biology

#### 2.1.1 Sample collection

Blood samples from 6 healthy donors were obtained from Research Blood Components. A signed written consent was obtained from all of the participants. All individuals recruited in this study were healthy Caucasian males between the ages of 21 and 55 years old. We decided to focus on only one sex to avoid the potentially confounding effects of sex-specific differences in gene expression level on response phenotypes (Ranz *et al.*, 2003; Rinn and Snyder, 2005). Only individuals self-reported as currently healthy, not under medication, and with no history of diseases such as malaria, tuberculosis, cancer, or hepatitis were included in the study. In addition, each donor's blood was tested for standard blood-borne pathogens, and only samples negative for all of the pathogens tested were used.

#### 2.1.2 *Mycobacterium tuberculosis* preparation

*Mycobacterium tuberculosis* preparation was performed as described by Barreiro *et al.* (Barreiro *et al.*, 2012). We infected dendritic cells (DCs) with a *Mycobacterium tuberculosis* (MTB) strain expressing green-fluorescent protein (H37Rv). This recombinant strain carries a pEGFP plasmid, which encodes a gene that confers resistance to hygromycin and harbors the GFP gene under the control of the mycobacterial Phsp60 constitutive promoter. Importantly, previous studies have shown (Kremer *et al.*, 1995; Tailleux *et al.*, 2003) that the presence of GFP in MTB does not alter growth or virulence of the bacilli under axenic conditions, relative to wild-type MTB. *M. tuberculosis* H37Rv was grown from a frozen stock to midlog phase in 7H9 medium (BD) supplemented with albumin-dextrose-catalase (ADC; Difco). We tested the virulence of the bacteria in the frozen stock by infecting C57BL/6 mice intranasally with 10<sup>3</sup> bacilli. After 21 and 42 d, we estimated a load of 10<sup>7</sup> bacteria in the mice lungs, indicating that the bacteria did not lose its natural virulence (Tanne *et al.*, 2009).

### 2.1.3 Isolation and infection of DCs

Isolation and infection of DCs were performed as described by Barreiro *et al.* (Barreiro *et al.*, 2012). Blood mononuclear cells from healthy volunteers were isolated by Ficoll-Paque centrifugation. Blood monocytes were purified from peripheral blood mononuclear cells by positive selection with magnetic CD14 MicroBeads (Miltenyi Biotech). Monocytes were then cultured for 5 days in RPMI 1640 (Invitrogen) supplemented with 10% heat-inactivated FCS (Dutscher), L-glutamine (Invitrogen), GM-CSF (20 ng/mL; Immunotools), and IL-4 (20 ng/mL; Immunotools). Cell cultures were fed every 2 days with complete medium supplemented with the cytokines previously mentioned. Before infection, we systematically checked the differentiation/activation status of the monocyte-derived DCs by flow cytometry, using antibodies against CD1a, CD14, CD83, and HLA-DR. All antibodies were purchased from Becton Dickinson. Only samples presenting the expected phenotype for non-activated DCs—CD1a<sup>+</sup>, CD14<sup>−</sup>, CD83<sup>−</sup>, and HLA-DR<sup>low</sup>—were used in downstream experiments. The resulting monocyte-derived DCs were then infected with MTB for 18 hours at a multiplicity of infection of 1-to-1. The choice of 18 hours is based on previous work, which revealed that the largest number of transcriptional changes following MTB infection could be captured at 18 h post-infection (Tailleux *et al.*, 2008).

### 2.1.4 MethylC-seq

The MethylC-seq protocol was performed as described by Lister *et al.* (Lister *et al.*, 2009). 1) Bisulfite Conversion. Starting material was 6 ug of DNA spiked with 3 ng of unmethylated Lambda DNA. Initial steps incorporated DNA fragmentation, end repair, addition of single A base, adaptor ligation, followed by bisulfite conversion via MethylCode™ Bisulfite Conversion Kit, Invitrogen. 2) Library Preparation. Illumina's TruSeq RNA Sample Preparation kit v2 was used for the preparation of the libraries. It involves purifying the poly-A containing mRNA molecules using oligo-dT attached magnetic beads. The cleaved RNA fragments are copied into first strand cDNA using reverse transcriptase and random primers. Second strand cDNA synthesis follows, using DNA Polymerase I and RNase H. The cDNA fragments then go through an end repair process, the addition of a single "A" base, and then



ligation of the adapters. The products are then purified and enriched with PCR. Libraries were run on the Illumina HiSeq 2000.

### **2.1.5 RNA-seq**

mRNA-Seq libraries were generated from total RNA with polyA<sup>+</sup> selection of mRNA using the TruSeq RNA Sample Prep Kit v2 (Illumina, San Diego, CA). Libraries were run on the Illumina HiSeq 2000.

Total RNA was isolated using the miRNeasy kit (Qiagen). Strand-specific libraries were prepared using the Illumina Total Stranded RNA Library kit, with the Ribo-Zero Gold, using the method specified by the manufacturer. Libraries were run on the Illumina HiSeq 2500.

### **2.1.6 ChIP-seq**

Samples were crosslinked with 1% w/v formaldehyde for 10 min at room temperature and immediately quenched for 5 min with 125mM Glycine at RT. The samples were then sonicated using the Bioruptors (Diagenode) and then ChIP DNA prepared using the IP-Star (Diagenode). Approximately 1 million cells were used for each ChIP and ~50,000 cells for the Input. The following antibodies were used: H3K4me1 (CST, 5326P, 1), H3K4me3 (CST, 9751BC, 7), H3K9me3 (MABI, 0318, 13001), H3K27me3 (MABI, 0323, 13001), H3K27ac (Abcam, Ab4729, GR119051), and H3K36me3 (MABI, 0333, 12003). ChIP and Input libraries were prepared using the Illumina Truseq Nano DNA kit, with alterations including: PCR enrichment (14 cycles) prior to size selection and utilizing the PippinPrep (SAGE Science) instead of the SPRI method for the size selection (200-400 bp). Libraries were run on the Illumina HiSeq 2500.

### **2.1.7 ATAC-seq**

The ATAC-seq protocol was performed as described by Buenrostro *et al.* (Buenrostro *et al.*, 2013) starting with 100,000 cells. Cells were lysed using cold lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl<sub>2</sub> and 0.05% IGEPAL CA-630). Immediately after lysis, nuclei were spun at 500g for 10 min using a refrigerated centrifuge. Immediately following the nuclei prep, the pellet was resuspended in the transposase reaction mix (25  $\mu$ L 2 $\times$  TD buffer,

2.5 µL transposase (Illumina) and 22.5 µL nuclease-free water) and let to incubate at 37 °C for 30 minutes. Directly following transposition the sample was purified using a Qiagen MinElute kit. Following purification, we amplified library fragments using 25 uL of 2X NEBnext PCR master mix, 0.3 uL of 100X SYBR Green I, 2.5 uL of each Nextera primers index 1 (i7) and 2 (i5), 10 uL of the transposed DNA, in a final volume of 50 uL using the PCR conditions previously described (Buenrostro *et al.*, 2013). To reduce GC and size bias in our PCR, we monitored the PCR reaction using qPCR in order to stop amplification before saturation (Buenrostro *et al.*, 2013). The libraries were purified using a Qiagen MinElute kit in 20 µL. Quality of the libraries was verified on a polyacrylamide gel and Bioanalyser but final quantification was done using KAPA Library Quant kit for Illumina Sequencing Platforms. Libraries were run on the Illumina Hiseq 2500.

## **2.2 Bioinformatics analyses**

### **2.2.1 Bioinformatics software**

Bedtools (Quinlan and Hall, 2010)

bismark (Krueger and Andrews, 2011)

Bowtie2 (Langmead and Salzberg, 2012)

Bsseq (Hansen *et al.*, 2012)

CENTPEDE (Pique-Regi *et al.*, 2011)

chromHMM (Ernst and Kellis, 2012)

DEseq2 (Anders *et al.*, 2013)

FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>)

GOrilla (Eden *et al.*, 2009)

HT-seq (<http://www-huber.embl.de/users/anders/HTSeq/doc/overview.html>)

Illumina Casava (version 1.8)

Macs2 (Zhang *et al.*, 2008)

ngs.plot (Shen *et al.*, 2014)

Picard (<http://picard.sourceforge.net/>)

R (<http://www.r-project.org/>)

Samtools (Li *et al.*, 2009)

Tophat2 (Kim *et al.*, 2013)

Trim\_galore ([http://www.bioinformatics.babraham.ac.uk/projects/trim\\_galore/](http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/))

### **2.2.2 MethylC-seq processing**

MethylC-seq libraries were sequenced using the Illumina HiSeq 2000, following the manufacturer's instructions. The standard Illumina pipeline was used for image analysis and base calling. We used the tool trim\_galore to trim off adapter sequence incorporated in the read and remove bases with a base quality score of 20. The resulting reads were mapped to human reference genome (GRCh37/hg19) and lambda phage genome using bismark (with the options --bowtie2 -p 12 -N 1), which uses a bisulfite converted reference genome (C-to-T and a G-to-A) for read mapping. Only reads that had a unique alignment and a minimum number of mismatch of one were retained. We removed PCR duplicates from each library and then merged the two libraries for the same sample. To obtain methylation evidence for each reference CpG, the number of reads for each site that were methylated and coverage which refers to the number of reads covering that position were evaluated using bismark methylation extractor. The context of each C was determined and each C was classified as CpG, CHH, or CHG, where H is either A, T, or C nucleotide. Methylation was estimated for each CpG locus using the number of C (methylated) and T (unmethylated) reads ( $C/C+T$ ).

### **2.2.3 ChIP-seq and ATAC-seq processing**

ChIP-seq and ATAC-seq libraries were sequenced using the Illumina HiSeq 2000, following the manufacturer's instructions. The standard Illumina pipeline was used for image analysis and base calling. We used the tool trim\_galore to trim off adapter sequence incorporated in the read and remove bases with a base quality score of 20. The resulting reads were mapped to human reference genome (GRCh37/hg19). The alignment software Bowtie 2 was used with the following options: -p 12 -N 1. Only reads that had a unique alignment and a minimum number of mismatch of one were retained. We removed PCR duplicates from each library and then merged the two libraries for the same sample.

#### **2.2.4 RNA-seq processing**

Non-strand-specific RNA-seq libraries from polyA<sup>+</sup> fraction were sequenced using the Illumina HiSeq 2000, following the manufacturer's instructions. The standard Illumina pipeline was used for image analysis and base calling. mRNA sequence reads were aligned to the human genome reference (GRCh37/hg19) sequence using the TopHat2 software package with a TopHat transcript index from RefSeq. We did not removed PCR duplicates from each library since these may represent true signal. Strand-specific libraries were processed similarly.

#### **2.2.5 Genome annotation**

Genomic regions were defined based on GRCh37/hg19 coordinates obtained from GREAT which uses only a subset of UCSC Known Genes (<http://bejerano.stanford.edu/help/display/GREAT/Genes/>). We limited our analysis to protein-coding genes. Since genes can have multiple RefSeq transcripts, we selected the leftmost transcription start site or rightmost transcription end site to associate a single transcript per gene in the positive strand. The reverse was done for genes on the negative strand. We defined promoters as regions 2.5 kb upstream of the TSS and gene bodies as the transcribed regions from the start to the end of transcription sites for each gene. Furthermore, exons are defined as regions from their specific coding start and end. Introns are defined as regions from the coding end of an exon to the coding start of the next exon.

#### **2.2.6 DC methylation profile**

We merged CpG methylation evidence from both strands and pooled the methylation data from both infected and non-infected samples to get a better estimate of methylation levels. We divided each genomic regions (upstream, gene body, downstream) in 20 bins from 5' to 3' end of protein-coding genes, and the mean mC/C level within each bin where individual CpG sites belong to was determined. Genes on the negative strand were reversed. To obtain mean methylation profiles for each genomic region oriented from 5' to 3'.

### **2.2.7 Identification of MTB-DMRs**

The summarized methylation estimates of strand-merged CpG sites were used to identify differences in methylation between infected and non-infected samples using the R package `bsseq`. `bsseq` implements a smoothing method that uses a local likelihood approach to estimate the smoothed probability of methylation at each site, taking into account the spatial correlation between nearby sites and placing greater weight on sites with higher coverage. We limited our analysis to CpG sites with a read depth  $\geq 4$  in at least 3 infected and 3 non-infected samples. We identified DMRs as regions containing at least 3 consecutive CpG sites that are significantly differentially methylated using a paired t-test ( $|t\text{-statistic}| > 3.9$  at  $P = 0.01$ ) and exhibited at least a 10% difference in mean methylation levels between treated and untreated samples.

### **2.2.8 Gene ontology (GO) analysis**

We used `GORilla` to test for enrichment of functional annotations among genes nearby MTB-DMRs, using all GREAT genes as a background set. Analysis was done with default parameters and results corrected for multiple testing by the method of Benjamini and Hochberg to control the false discovery rate.

### **2.2.9 ChIP-seq data analysis**

We used the `MACS2` (model-based analysis of ChIP-Seq) peak-finding algorithm to identify regions of ChIP-Seq enrichment over background (broad mode). A false discovery rate threshold of enrichment of 0.05 was used for all data sets. The resulting genomic coordinates in bed format were further used in `ChromHMM` for chromatin annotation. The following parameters were used: `-Xmx20g -jar ChromHMM.jar BinarizeBed hg19 -Xmx2000M -jar ChromHMM.jar LearnModel 10 hg19`.

### **2.2.10 Footprinting analysis**

The input for `CENTIPEDE` included the PWM score, conservation (PhyloP) and ATAC-seq counts within  $\pm 100$  bp of each genomic region matching a motif.

### **2.2.11 Identifying genes differentially expressed after MTB infection**

A python script called HTSeq counts how many reads map to each feature using the output BAM file and a GTF file containing genomic structures or features. Non-expressed genes (read count = 0 in at least 3 infected or 3 non-infected samples) were discarded. Using the normalized counts for each gene, we identified genes whose expression levels were significantly altered following MTB infection of DCs using the Bioconductor DESeq2 with cutoffs: absolute log2 fold change > 1 and 1% false discovery rate (FDR).

### 3 Results

---

**Alain Pacis**, Ludovic Tailleaux, John Lambourne, Vania Yotova, Anne Dumaine, Francesca Luca, Jean-Christophe Grenier, Kasper Hansen, Maio Yu, Chuan He, Tomi Pastinen, Roger Pique-Regi, Yoav Gilad, Luis B Barreiro. Dynamic epigenetic changes in immune responses to infection in human dendritic cells.

---

*Declaration of contribution to “Dynamic epigenetic changes in immune responses to infection in human dendritic cells”: As primary author of this manuscript, I developed the entire bioinformatics pipeline to process and analyze the sequencing data. I also wrote the draft of the manuscript including figures, figure legends, supplementary information and experimental approach section. Ludovic Tailleaux performed the infection of dendritic cells with Mycobacterium tuberculosis. John Lambourne and Tomi Pastinen performed chromatin immunoprecipitation of histone marks and constructed ChIP-seq libraries. Vania Yotova constructed bisulfite-treated DNA libraries and RNA-seq libraries. Anne Dumaine constructed ATAC-seq libraries. Jean-Christophe Grenier provided assistance in the bioinformatics analysis. Kasper Hansen provided the bsseq program to identify differentially methylated regions. Maio Yu and Chuan He performed the protocol for TAB-seq to obtain 5hmC data (not shown used in this thesis). Roger Pique-Regi and Francesca Luca helped with the footprinting analysis. Yoav Gilad provided with reagents and analytical tools. Luis B Barreiro designed and supervised the study.*

**Dynamic epigenetic changes in immune responses to infection  
in human dendritic cells**

Alain Pacis<sup>1,2</sup>, Ludovic Tailleux<sup>3</sup>, John Lambourne<sup>4</sup>, Vania Yotova<sup>1</sup>, Anne Dumaine<sup>1</sup>,  
Francesca Luca<sup>5</sup>, Jean-Christophe Grenier<sup>1</sup>, Kasper Hansen<sup>6</sup>, Maio Yu<sup>7</sup>, Chuan He<sup>7</sup>, Tomi  
Pastinen<sup>4</sup>, Roger Pique-Regi<sup>5</sup>, Yoav Gilad<sup>8</sup>, Luis B Barreiro<sup>1,9,\*</sup>

<sup>1</sup>CHU Sainte-Justine Research Center, Department of Genetics, Montreal, H3T1C5, Canada.

<sup>2</sup>University of Montreal, Department of Biochemistry, Montreal, H3T1J4, Canada. <sup>3</sup>Institut Pasteur, Mycobacterial Genetics Unit, Paris, 75015, France. <sup>4</sup>Génome Québec Innovation Centre, Department of Human Genetics, McGill University, Montréal, H3A0G1, Canada.

<sup>5</sup>Wayne State University, Center for Molecular Medicine and Genetics, Detroit, MI, 48202.

<sup>6</sup>Johns Hopkins Bloomberg School of Public Health, Department of Biostatistics, Baltimore, MD, 21205. <sup>7</sup>University of Chicago, Department of Chemistry and Institute for Biophysical Dynamics, Chicago, IL, 60637. <sup>8</sup>University of Chicago, Department of Human Genetics, Chicago, IL, 60637. <sup>9</sup>University of Montreal, Department of Pediatrics, Montreal, H3T1J4, Canada.



### 3.1 SUMMARY

DNA methylation is an important epigenetic mark in mammals. Although methylation at the 5' position of cytosine (5mC) is recognized as a stable epigenetic modification, it is becoming increasingly viewed as a more dynamic process which involves both active methylation and demethylation pathways. While the dynamics of DNA methylation has been well-characterized in mammalian development and normal cellular function, little is known about its regulatory implications in immune responses. To that end, we performed comprehensive transcriptional and epigenetic profiling of primary dendritic cell (DC) samples from humans, before and after infection with *Mycobacterium tuberculosis* (MTB). Our results provide the first complete genomic portrait of the extensive epigenetic remodeling occurring in primary DCs in response to a bacterial infection. We found that active changes in DNA methylation are pervasive, identifying 3,926 MTB-induced differentially methylated regions (MTB-DMRs). MTB-DMRs show a striking overlap with genomic regions marked by histones associated with enhancer activity. ATAC-seq footprinting analysis revealed that regions that change methylation were actively bound by immune-related TFs prior to MTB-infection suggesting that these domains are likely to represent enhancer elements in a poised state. Finally, we found that differentially expressed genes are associated with MTB-DMRs which indicates a potential mechanistic link between DNA methylation and gene expression. Our data suggests that active changes in methylation play an essential and previously unappreciated role at controlling of the regulatory programs engaged by DCs in response to a bacterial infection.

### 3.2 INTRODUCTION

An effective immune response requires the coordinated action of numerous cell types in the immune system. Dendritic cells (DCs) are innate immune cells present virtually everywhere in the body and are in front line of defense against invading microbial agents and other insults (Banchereau and Steinman, 1998). DCs typically sense immune challenge through pattern recognition receptors, such as Toll-like receptors (TLRs), located on the cell surface or on the endosomal/lysosomal compartment. TLRs recognize molecular signatures that are broadly shared by pathogens, collectively referred to as pathogen-associated molecular patterns (PAMPs) (Medzhitov, 2001; Janeway and Medzhitov, 2002; Medzhitov, 2007). Interaction of TLRs with foreign substances activates DCs, leading to their migration to the draining lymph nodes where they present antigens to naive T cells and initiate adaptive immunity (Banchereau and Steinman, 1998; Cooper, 2009).

Upon stimulation, DCs undergo pervasive changes in their transcriptional program, mobilizing hundreds of genes involved in immune-related processes (i.e., antimicrobial peptides, reactive oxygen species, adhesion molecules, cell death, chemokines and cytokines) in a rapid and highly choreographed fashion (Smale, 2010; Smale *et al.*, 2014). This is achieved with the help of signal-dependent transcription factors (TFs), including NF- $\kappa$ B, STAT, AP-1, and interferon regulatory factors (IRFs), which bind to gene regulatory regions of the genome where they function to initiate recruitment of various co-activators (Medzhitov and Horng, 2009; Smale, 2010, 2011). Epigenetic processes, such as histone modifications and DNA methylation, are recognized as important permissive or suppressive factors playing an important role in modulating the access of TFs to *cis*-acting DNA regulatory elements by regulating chromatin dynamics and therefore have a significant impact in gene expression. Recently, significant progress has been made in our understanding of epigenetic control of many immune processes, such as in hematopoietic cell development and lineage decisions (Broske *et al.*, 2009; Klug *et al.*, 2010; Hodges *et al.*, 2011; Hogart *et al.*, 2012; Schlesinger *et al.*, 2013). Similarly, immune responses are subject to epigenetic control to ensure its proper function. Using recent epigenomic profiling technologies, a number of studies have highlighted the paramount importance of epigenetic mechanisms in the regulation of complex gene

expression programs that underlie immune responses (Foster *et al.*, 2007; Barish *et al.*, 2010; Ghisletti *et al.*, 2010; Garber *et al.*, 2012; Kaikkonen *et al.*, 2013; Ostuni *et al.*, 2013). However, these studies have been focused on dynamics of histone marks and have been primarily performed in innate immune cells in mouse models. No attempts have been made thus far to map the DC epigenome, particularly patterns of DNA methylation, in humans. More importantly, little is still known about the extent of epigenetic remodeling in DCs in response to environmental stress.

The immune system has been a powerful and effective system to study how the epigenome impacts gene expression during immune responses. Similar to macrophages, activated DCs can serve as a paradigm for dissecting environmentally responsive transcriptional control in an experimentally tractable system. In this study, we investigated epigenetic changes induced by a pathogenic bacteria, *Mycobacterium tuberculosis* (MTB), the causative agent of tuberculosis (TB) in humans. We chose to work with MTB because it is a pathogen of major importance to human health and for which the immunological interactions with DCs have been extensively studied (Barreiro *et al.*, 2012).

To investigate dynamic epigenetic patterns and transcriptional outputs associated with bacterial infection, we infected monocyte-derived dendritic cells with a virulent strain of MTB for 18 hours. Following infection, we measured both genome-wide methylation levels and expression in non-infected and infected DCs using whole-genome bisulfite sequencing (MethylC-seq) and RNA sequencing (RNA-seq), respectively. To complement this data, DC samples were subjected to chromatin immunoprecipitation sequencing (ChIP-seq) for six histone marks (H3K4me1, H3K4me3, H3K27me3, H3K27ac, H3K36me3, and H3K9me3) and assay for transposase-accessible chromatin using sequencing (ATAC-seq) for open chromatin profiling. Our study provides unprecedented insight into the transcriptional and epigenetic dynamics during bacterial infection in primary dendritic cells. It also offers a unique opportunity to explore the epigenetic mechanisms underlying the regulation of immune responses against infectious diseases.

### 3.3 RESULTS

#### Single-base Resolution Mapping of DC Methylome

We constructed directional bisulfite-treated libraries using standard Illumina protocol for each DC sample in both conditions and generated a total of ~8 billion single-end reads by next-generation sequencing. Seventy-four percent of the reads were uniquely aligned to either strand of the reference genome (hg19), achieving a combined average coverage per CpG site of ~10× for each sample. We detected approximately 28 million CpG sites in each sample, which covers approximately 90% of total CpG sites. Comparison of methylcytosines of biological replicates in both conditions revealed a high concordance of CpG methylation status (Pearson  $r$  of 0.8-0.9; **Table S1**). Overall, DNA methylation in DCs was found almost exclusively in the CpG context with 76% of CpG sites being highly methylated (0.8-1.0 methylation) and only less than 1% CpH methylation, consistent with data from other fully differentiated cell types (**Figure S2A**). In line with other reports, methylation is high ubiquitously except near transcription start sites (TSS) and CpG islands (**Figures S2B, S2C**) (Lister *et al.*, 2009; Stadler *et al.*, 2011). As expected, hierarchical clustering analysis of the MethylC-seq data, which included human hematopoietic, embryonic, fibroblast and prefrontal cortex cells for comparison, revealed that the DC methylome was more similar to that of hematopoietic cells (**Figure S2D**).

#### MTB Infection Induces Genome-wide Active Changes in DC Methylome

To explore the scope of dynamic DNA methylation during MTB infection, we looked at differences in CpG methylation between untreated and MTB-infected DCs in a genome-wide scale. We utilized a set of CpG sites with  $\geq 4$  sequence coverage in at least half of the DC samples for both conditions, sufficient to interrogate ~20 million CpGs. We defined MTB-induced differentially methylated regions (MTB-DMRs) as regions of 3 or more adjacent CpG sites exhibiting a minimal mean difference of CpG methylation level of 0.1 between the two groups ( $p < 0.01$ ; see **Experimental Procedures**). Overall, we identified 15,904 dynamic CpGs distributed across 3,926 discrete DMRs (1,888 hyper- and 2,038 hypomethylated; **Figure 1A**). Interestingly, most of the MTB-DMRs are small with the majority of the regions

comprising of 3 dynamic CpGs (**Figure S3A**). Hypermethylated DMRs have a mean length of 137 bp, significantly shorter than for hypomethylated regions (182 bp,  $p = 6.27 \times 10^{-18}$ , Wilcoxon) (**Figure 1B**). We then mapped the MTB-DMRs to nearby gene annotation to assess the genomic distribution of changes in DNA methylation. We associated a total of 1,358 and 1,423 unique genes to hyper- and hypomethylated regions respectively. Interestingly, the vast majority of MTB-DMRs are not located near promoters (7% at -2.5 kb from transcription start sites (TSSs)). Rather, these epigenetically dynamic regions are predominantly found several kilobases away from TSSs, at CpG-poor intragenic (49%) and intergenic regions (44%) (**Figures 1C, 1D and 1E**). Collectively, these data show that MTB infection induces extensive changes in DNA methylation in fully differentiated DCs and that such changes primarily occur at putative distal regulatory elements.

### **MTB-DMRs Have an Enhancer-Associated Chromatin Signature**

To gain a more complete picture of the underlying epigenetic dynamics and to investigate the regulatory events during MTB infection, we next collected ChIP-seq data for 6 histone marks (H3K4me1, H3K4me3, H3K27ac, H3K27me3, H3K36me3 and H3K9me3) in infected and non-infected DCs. Recent advances in epigenomic profiling technologies have allowed the identification of chromatin features associated to enhancers. These epigenomic signatures include high enrichment of monomethylated K4 on the histone H3 N-terminal tail (H3K4me1) (Heintzman *et al.*, 2009) and low enrichment of H3K4me3 compared to promoters. We examined the relative enrichments of each histone marks at MTB-DMRs and found that proximal and distal DMRs are enriched for H3K4me3<sup>hi</sup>/H3K4me1<sup>lo</sup> and H3K4me1<sup>hi</sup>/H3K4me3<sup>lo</sup> marks, respectively (**Figure 2A**) suggesting that MTB-DMRs are acting as enhancer elements. To further understand the functional significance of MTB-DMRs, we generated gene regulatory annotation maps from the panel of ChIP-seq data using the hidden Markov model-based ChromHMM chromatin segmentation program (Consortium *et al.*, 2012; Ernst and Kellis, 2012). Consistent with the Refseq-based annotation, there are very few MTB-DMRs localized to ChromHMM-annotated promoter regions (state 1). An important insight enabling the distinction of active enhancers from inactive/poised enhancer elements containing H3K4me1, is through the presence (state 2) or absence (state 3) of histone

acetylation (H3K27ac). Interestingly, we found a third type of enhancers which exhibit the following specific combination of histone marks: H3K4me1+H3K27ac+H3K4me3 (state 4). We found that 41% of all the MTB-DMRs overlap with a putative enhancer region in non-infected DCs, a 6.4-fold enrichment altogether as compared to genome-wide expectations ( $p < 2.2 \times 10^{-16}$ ; **Figures 2B, 2C**). Interestingly, the overlap with active enhancer regions, particularly at state 2, becomes even stronger when chromatin states are defined using MTB-infected DCs (**Figure 2C**).

Previous studies of developmentally dynamic enhancers have shown that enhancer activation is associated with mCG hypomethylation of the enhancer, whereas enhancer inactivation is associated with enhancer hypermethylation (Gifford *et al.*, 2013; Lister *et al.*, 2013). We therefore examined chromatin transition states at hyper- and hypomethylated regions separately. Regions that show increased CG methylation upon MTB infection does not seem to be associated with specific local chromatin modifications (72% of hyper-DMRs). Conversely, decrease in methylation resulted in significant changes in chromatin at pre-existing enhancers, particularly gaining either H3K27ac ( $4 \rightarrow 3$ ) or H3K4me3 marks ( $3 \rightarrow 2$ ). At the same time, hypomethylated regions give rise to *de novo* enhancer regions, represented by transitions  $8 \rightarrow 3$  and  $8 \rightarrow 4$  (**Figure 2D**). Consistent with the transition states above, patterns of histone patterns in distal regions that are hypermethylated upon infection are identical to those observed in unstimulated conditions (**Figure S4**), indicating that hypermethylation may not necessarily indicate inactivation in these genomic regions, at least in a non-developmental context. Also, hypomethylated regions generally gain localized histone modifications characteristic of enhancers (H3K4me1 and H3K27ac) as well as H3K4me3 (**Figures 2E**). Interestingly, the central sequences within hypomethylated regions are depleted of these three histone marks, which may indicate selective nucleosome eviction thorough binding of transcription factors and co-activators such as p300 (Ghisletti *et al.*, 2010) after MTB infection.

A recent genome-wide study in mouse cortical neurons have demonstrated that neuronal activity-regulated enhancers recruit RNA polymerase II (RNAPII) and transcribes bi-directionally a novel class of enhancer RNAs (eRNAs) (Kim *et al.*, 2010). Since hypomethylated regions show characteristics of transcriptional activation in an activity-

dependent manner, this raises the possibility that RNA transcription may occur at these enhancer regions. To test that hypothesis, we sequenced total RNA obtained from infected and non-infected DCs after ribosomal RNA was depleted. To distinguish possible enhancer-derived transcripts from mRNA transcripts, we sought evidence of RNA transcription specifically at 833 hypomethylated regions located outside of annotated genes. We observed dynamic changes in the levels of these enhancer RNAs (eRNAs) upon infection (**Figure 2F**) and synthesis of eRNAs seems to initiate near enhancer centers where perhaps signal-dependent transcription factors and RNAPII are bound and to proceed bi-directionally. These observations indicate that hypomethylated regions are not merely sites where transcription factors bind and recruit RNAPII that might subsequently be delivered to promoters, but are also sites where RNA synthesis occurs in response to MTB infection.

### **MTB-DMRs are Bound by Signal-Dependent Transcription Factors**

Gene regulatory regions are usually characterized by DNase I hypersensitivity (Thurman *et al.*, 2012; John *et al.*, 2013) or transposase (Tn5) accessibility (Buenrostro *et al.*, 2013) as DNA is usually histone-free in gene regulatory regions so transcription factors can bind to these regions. To explore the potential regulators of chromatin dynamics during MTB infection, we assayed for transposase-accessible chromatin using sequencing (ATAC-seq) (Buenrostro *et al.*, 2013). Tn5-accessibility profile in hypermethylated regions is roughly similar after MTB infection, consistent with the observed stable chromatin features. In contrast, we observed markedly higher Tn5 accessibility densities among hypomethylated regions commensurate with the gain of localized histone marks (**Figure 3A**). This is illustrated in three different regions that lose methylation within the genebody of *NFKB1*. Here all three hypo-DMRs are located in intronic regions are classified as enhancers by chromHMM-based gene regulatory region annotation and lead to active chromatin conformation (**Figure 3B**).

If MTB-DMRs correspond to enhancers controlling inflammatory gene expression, an obvious prediction is that they should be enriched for binding sites for stimulus-activated TFs required for the deployment of the inflammatory gene expression program, like NF- $\kappa$ B (Rel) and interferon regulatory factors (IRF) family members (Taniguchi *et al.*, 2001; Hayden and Ghosh, 2008). Bound transcription factors to genomic DNA protect the underlying binding

motifs from DNase I cleavage or tn5 transposase accessibility and thereby, creating “genomic footprints” (Neph *et al.*, 2012; Buenrostro *et al.*, 2013) that can be used to reliably predict the binding of hundreds of TFs. We inferred ~5.8 million footprints from ATAC-seq data for both conditions using *CENTIPED*E program (Pique-Regi *et al.*, 2011). Hypermethylated regions are enriched in footprints containing motifs for IRF and other TFs potentially implicated in different roles in immune responses. Once again, since these regions do not exhibit drastic changes in their chromatin accessibility, we did not expect any change in TF binding either and indeed found so (**Figure 3C**). Hypomethylated regions in particular are more enriched in footprints tied to motifs for NF- $\kappa$ B, which showed the most significant enrichment ( $p < 2.2 \times 10^{-16}$ ), in addition to IRF. Strikingly, transcription factors NF- $\kappa$ B or IRF, as well as others, do not globally show any significant evidence for increased binding after infection (**Figure 3C**).

### **MTB-DMRs are Linked to Changes in Expression of Inflammatory Genes**

To gain insight on the global transcriptional dynamics during MTB infection, we performed RNA-seq on poly-A fractions on the same samples for which we collected whole-genome methylation profiles. Hierarchical clustering shows that infected samples forms a distinct cluster from non-infected samples based on their changes in gene expression after MTB infection (**Figure S5**). Using a difference in normalized expression level higher than absolute 1-fold as a cutoff, we identified 3,637 protein-coding genes that were significantly differentially expressed (DE; FDR < 0.01). We next asked whether dynamic DNA methylation is implicated in MTB-induced changes in gene expression. To limit arbitrariness in gene assignment, we considered only MTB-DMRs that were within  $\pm 100$  kb from the closest TSS. In general, MTB-DMRs are more likely to be associated with genes changing expression after infection (28% versus 21% that are expected by chance alone;  $p = 1.78 \times 10^{-12}$ , chi-squared test; **Fig. 4A**). Notably, we found that DE genes are more strongly linked to hypomethylated ( $p = 2.01 \times 10^{-13}$ , chi-squared test) than hypermethylated regions ( $p = 0.0015$ ) suggesting that loss of methylation has a more pronounced role in directly regulating immune-specific genes. Moreover, when compared to a control group consisting of all the DE genes, the set of DE genes associated with hypomethylated regions was enriched for up-regulated transcripts at high statistical significance ( $p = 1.88 \times 10^{-14}$ , chi-squared test; **Figure 4B**). Gene set enrichment



analysis further revealed that genes associated with hyper- and hypomethylated regions are more over-represented in functional categories relevant in negative and positive regulation of immune responses respectively, such as in signaling and regulation of response to stimulus (**Figure 4C**). Moreover, hyper-DMRs are associated with genes particularly involved in negative regulation of biosynthetic and metabolic processes. This strongly suggests that hyper-DMRs are implicated in down-regulating genes that are involved in negative regulation of multiple biological processes since methylation is generally known as an epigenetic mechanism for repression.

### 3.4 DISCUSSION

In this study, we have generated transcriptomic and epigenomic data in non-infected and MTB-infected primary dendritic cells by high-throughput sequencing. This allowed, for the first time, for assessment of how the methylation landscape of DCs is rapidly modified in response to bacterial infection. Our results show that infection of human DCs with MTB leads to a marked remodeling of its methylome, with several thousand regions showing significant changes in methylation levels upon infection. Interestingly, we show that MTB-DMRs are located in regions of low-CpG density far away from TSS and indeed, they show a striking overlap with enhancer regions defined by histone marks. Furthermore, MTB-DMRs are associated with a large fraction of the genes that are differentially expressed after MTB infection.

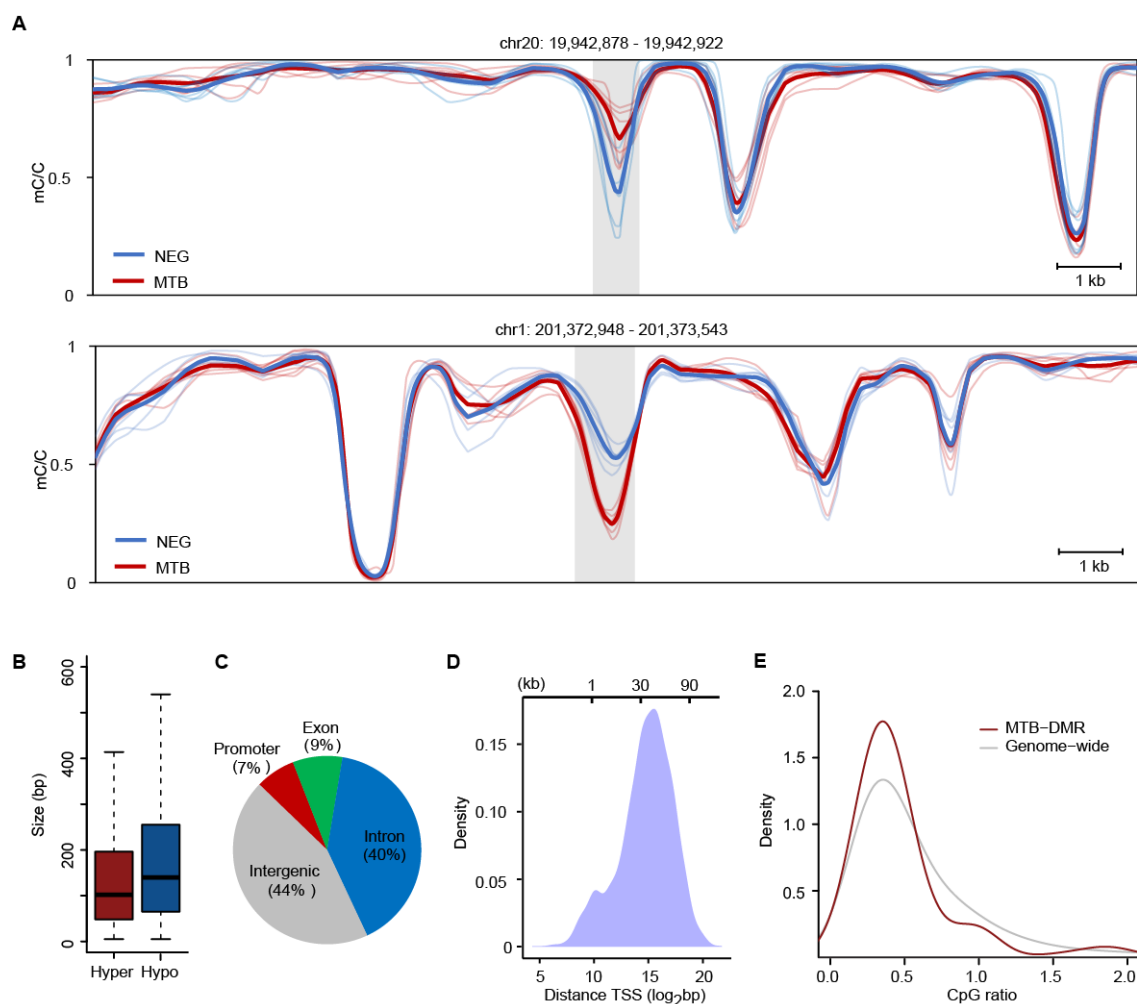
Footprinting analysis revealed that MTB-DMRs are bound by TFs known to play a key role in immune responses to infection, such as NF- $\kappa$ B and IRFs. Strikingly, our results further indicate that differential methylation did not result to any significant changes in the binding of these TFs. Therefore, it is highly conceivable that transcriptional regulators that interact indirectly with the DNA sequence are responsible for the observed changes in methylation. Indeed, it is known that many transcriptional regulators are positioned to interact indirectly with the DNA sequence of some target sites through mechanisms such as tethering (Biddie *et al.*, 2011; Neph *et al.*, 2012). Binding of such transcriptional regulators is likely to be missed since footprinting assays cannot detect protein-protein interactions. For instance, in hypomethylated regions, we found an increase in H3K27ac mark after MTB infection

suggesting the recruitment of p300 co-activator in a stimulus-regulated manner. However, footprinting assay did not detect any binding evidence for CREB-binding protein (CBP) or p300 acetyltransferases consistent with reports showing that CBP/p300 interacts with DNA-bound TFs (Ogryzko *et al.*, 1996; Martinez-Balbas *et al.*, 1998). The recruitment of co-activators as well as PolII might also contribute to the opening of the chromatin. Furthermore, the fact that TFs are already bound at hypo-DMRs prior to MTB infection strongly suggests that hypo-DMRs are maintained in an open chromatin conformation that is easily accessible for co-factor binding. This poised state potentially confers an effective and rapid response to DCs against MTB infection. Indeed, this is supported by the high levels of H3K4me1 in these regions at basal state. In addition, 5-hydroxymethylcytosine has been shown to contribute to the poised chromatin signature found at enhancers (Yu *et al.*, 2012). It will be interesting to examine whether 5hmC is enriched in hypo-DMRs prior to infection.

There is mounting evidence that innate immune responses in mammals have the capacity to exert immunological memory upon re-stimulation (Quintin *et al.*, 2014; Smale *et al.*, 2014). A recent study by Ostuni *et al.* shows that the stimulation of mouse macrophages with different stimuli (e.g., LPS, IFN $\gamma$ , IL-4) induces the appearance of a few thousand *de novo* putative enhancer regions, coined “latent enhancers”, that were absent in naïve cells (Ostuni *et al.*, 2013). More importantly, H3K4me1 marks were found to persist for days in latent enhancers when stimulation ceased, which potentially confers the ability to these regions to contribute to a faster and stronger transcriptional response to repeated challenges with the same stimulus. This study highlights the implication of epigenetic histone modifications in mediating innate immune memory. Although 5-methylcytosine has long been regarded as a stable marker, recent data, including the data reported herein, shows that 5mC can dynamically change in response to various environmental cues (Guo *et al.*, 2011). Indeed, this is supported by the recent discovery of Tet (ten-eleven-translocation) proteins as new epigenetic modifiers responsible for conversion of 5-methylcytosine to 5-hydroxymethylcytosine as part of the active demethylation process (Bhutani *et al.*, 2011). This opens the possibility that DNA methylation can serve as a mechanism for stimulus-induced epigenetic control and long-lasting plasticity in innate immune cells. Thus, in addition to histone modifications, epigenetic DNA modifications may expand the capacity of such cells to perpetuate transient stimuli into long-

lasting covalent modifications for future encounters with similar stimuli. Based on our current data, it is hard to identify whether changes in DNA methylation remained fixed overtime. Although it remains unclear whether transcription factor binding precedes changes in DNA methylation, it is hard to pinpoint at which time point during MTB infection is dynamic methylation associated with binding or unbinding of TFs and ultimately, to changes in gene expression. Further work is therefore required to investigate the implication of DNA methylation in innate short-term memory, ideally through time-course experiments. Also, it will be imperative to examine whether changes in DNA methylation could be maintained in the absence of a sustained stimulation.

These observations suggest that active and rapid changes in methylation might play a critical role in transcriptional regulation of an efficient immune response to infection. Our study also provides a comprehensive resource and offer a new framework for testing the role of the epigenome in immune responses. We expect that our findings will stimulate future work in unraveling the epigenetic mechanisms contributing to immune responses associated with other infectious diseases or immune challenges.



**Figure 1. DC Methylome is Dynamically Remodeled during MTB Infection**

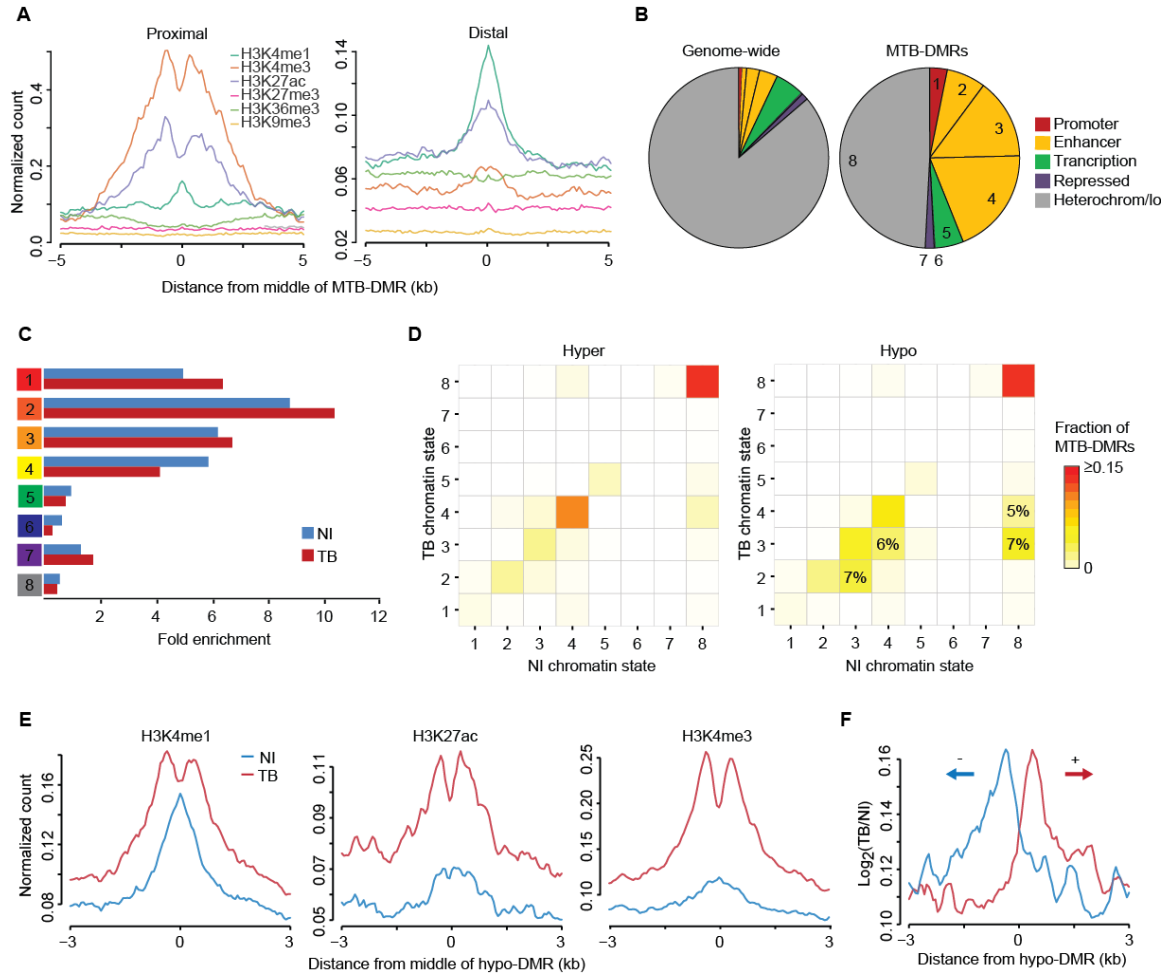
(A) Smoothed methylation values (y-axis) for 6 MTB-infected samples (red) and 6 non-infected samples (blue) showing an example of a hyper- (up) and a hypomethylated region (down; gray shading). Genomic coordinates of MTB-DMRs are indicated on top. Thick lines represent the average methylation for each condition.

(B) Boxplots comparing the distribution of genomic sizes in base pairs (bp) of hyper- and hypomethylated regions.

(C) Refseq annotation of MTB-DMRs.

(D) Distribution of distances (bp) of MTB-DMRs to nearest genes in  $\log_2$  (down) and in linear forms (up; in kilobases).

(E) Distributions of CpG densities of 250-bp windows flanking MTB-modified CpGs and all CpGs in the human genome.



**Figure 2. MTB-DMRs Have Chromatin Characteristics of Enhancers**

(A) Composite plots of patterns of histone marks, 5 kb around midpoints of proximal (left) and distal (right) MTB-DMRs.

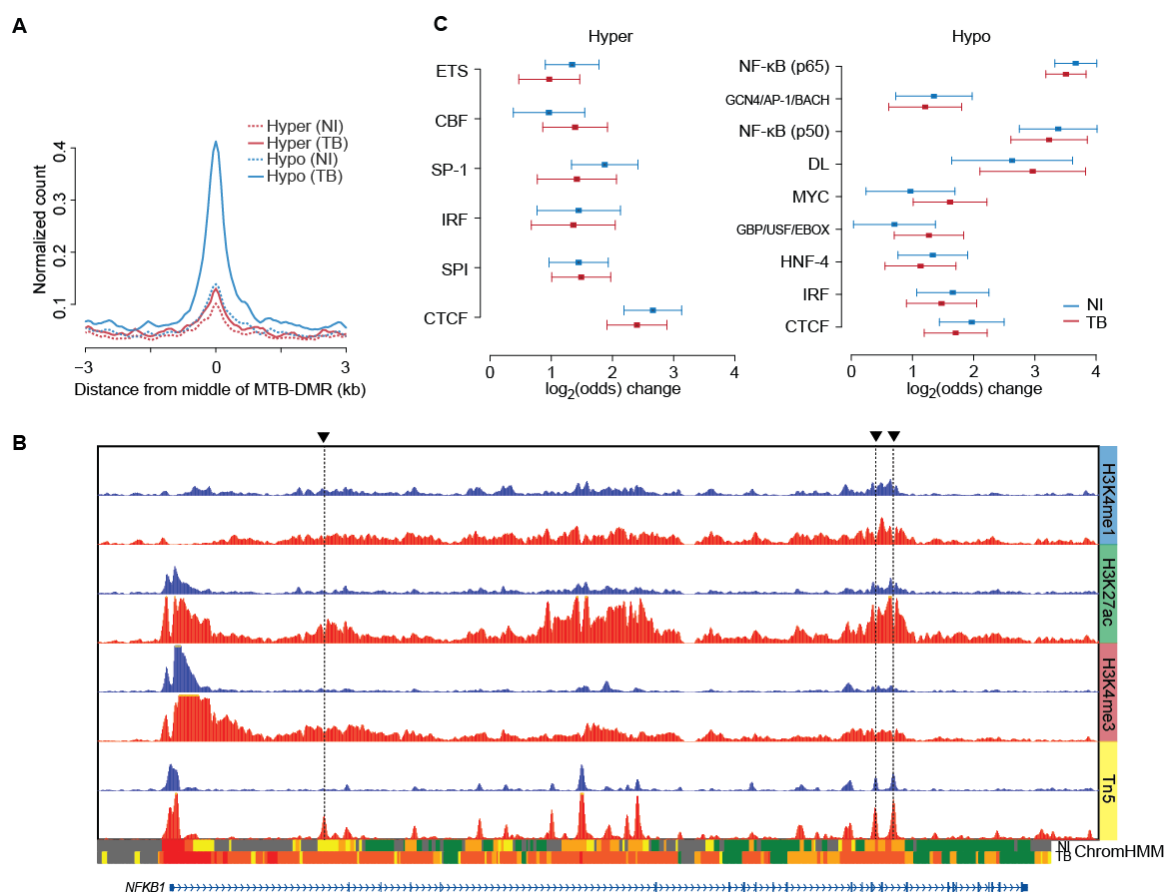
(B) Chromatin state annotation of MTB-DMRs using chromHMM. The number code annotation of the chromatin state map is as follows: 1) promoter, 2) strong enhancer (H3K4me1+H3K27ac+H3K4me3), 3) strong enhancer (H3K4me1+H3K27ac), 4) weak enhancer (H3K4me1), 5) transcription, 6) repressed state 1, 7) repressed state 2 and, 8) heterochromatin or no signal.

(C) Relative enrichment of ratio of MTB-DMRs in non-infected and MTB-infected states, compared with the representation of chromatin states genome-wide.

(D) Chromatin transition states at hyper- (left) and hypo-DMRs (right). The x- and y-axis represent chromatin states before and after infection respectively.

(E) Composite plots of patterns of histone marks before and after MTB infection, 3 kb around midpoints of hypo-DMRs.

(F) Profile of RNA expression based on ribo-minus RNA fractions at intergenic hypo-DMRs, in  $\log_2(\text{infected}/\text{non-infected})$ . (+) and (-) denote transcripts in forward and reverse strands respectively.

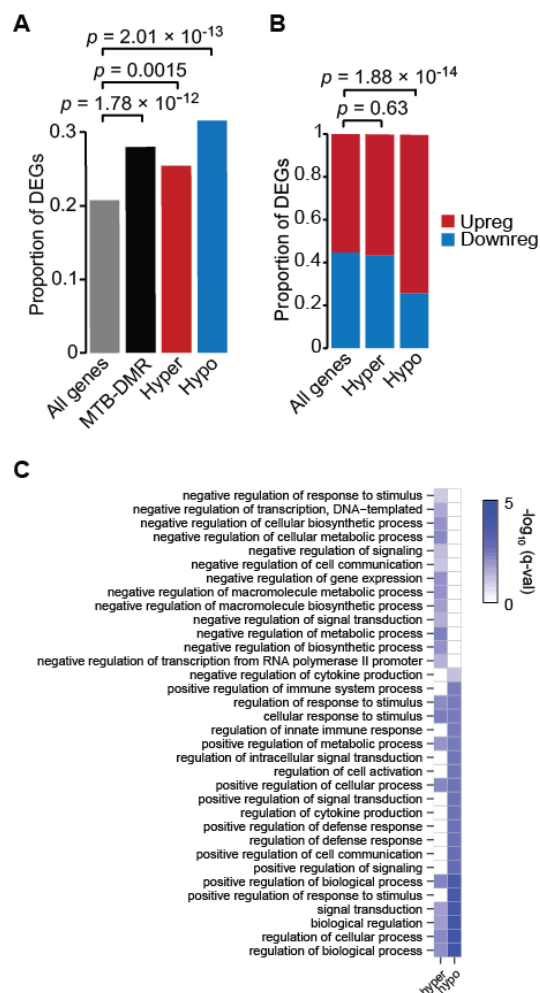


**Figure 3. MTB-DMRs are Enriched for Immune-Specific Transcription Factor Footprints**

(A) Tn5-accessibility profiles before and after MTB infection, 3 kb around midpoints of MTB-DMRs.

(B) Chromatin dynamics at 3 separate hypo-DMRs (indicated by arrows) within *NFKB1* genebody, shown by normalized ChIP-seq (H3K4me1, H3K27ac, and H3K4me3) and chromHMM tracks. ChIP-seq tracks in blue and red represent non-infected and MTB infected states respectively.

(C) Enrichment of transcription factor footprints associated with hyper- and hypo-DMRs, before and after MTB infection.



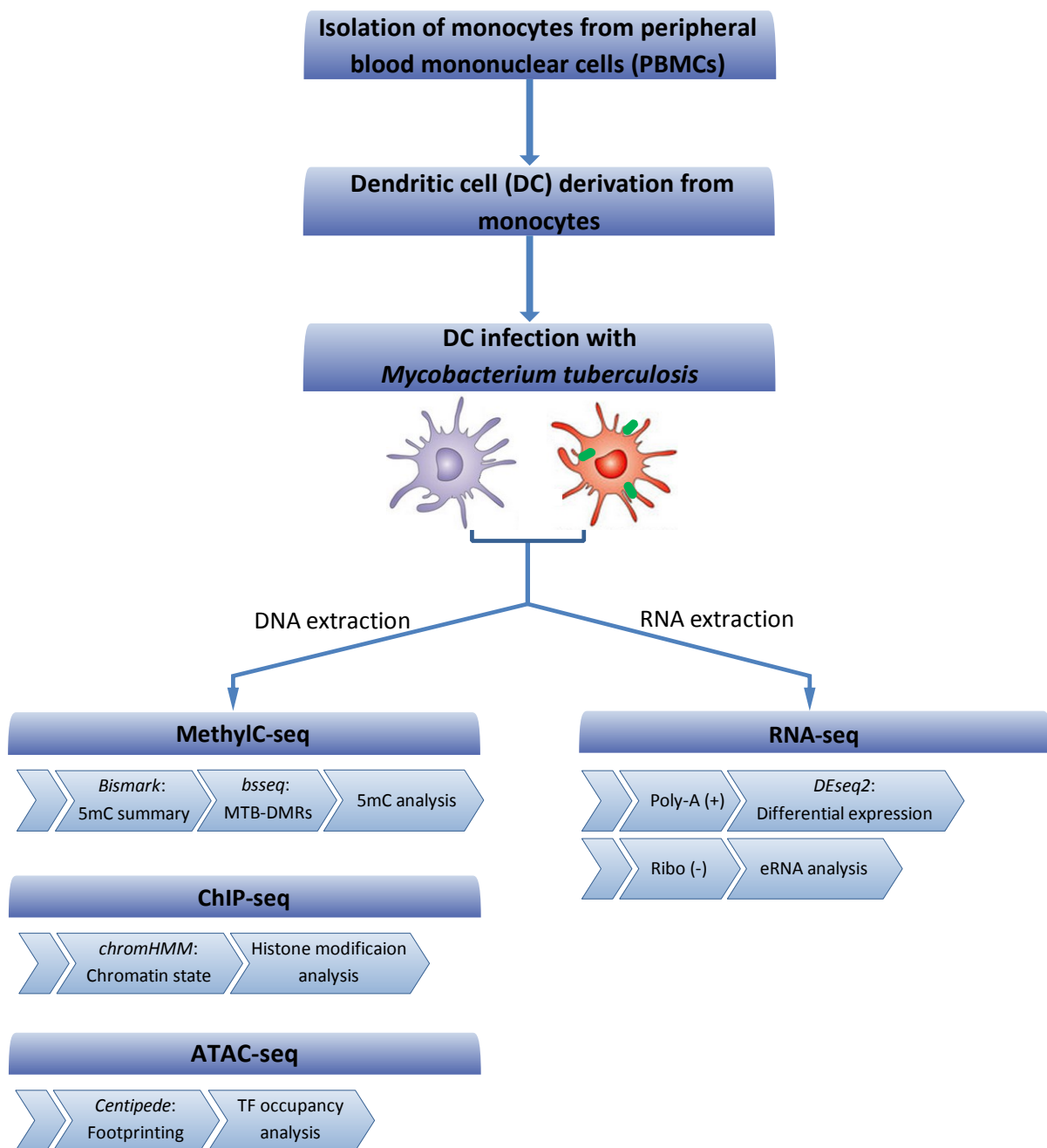
**Figure 4. Differential Methylation is Coupled to Differential Gene Expression**

(A) Proportion of differentially expressed genes (DEGs; y-axis) observed among all tested genes (control) and among genes associated to MTB-DMRs following infection with MTB.

(B) Proportion of up- and down-regulated genes associated with hyper and hypo-DMRs. (C)

GO categories associated with hyper- and hypo-DMRs.



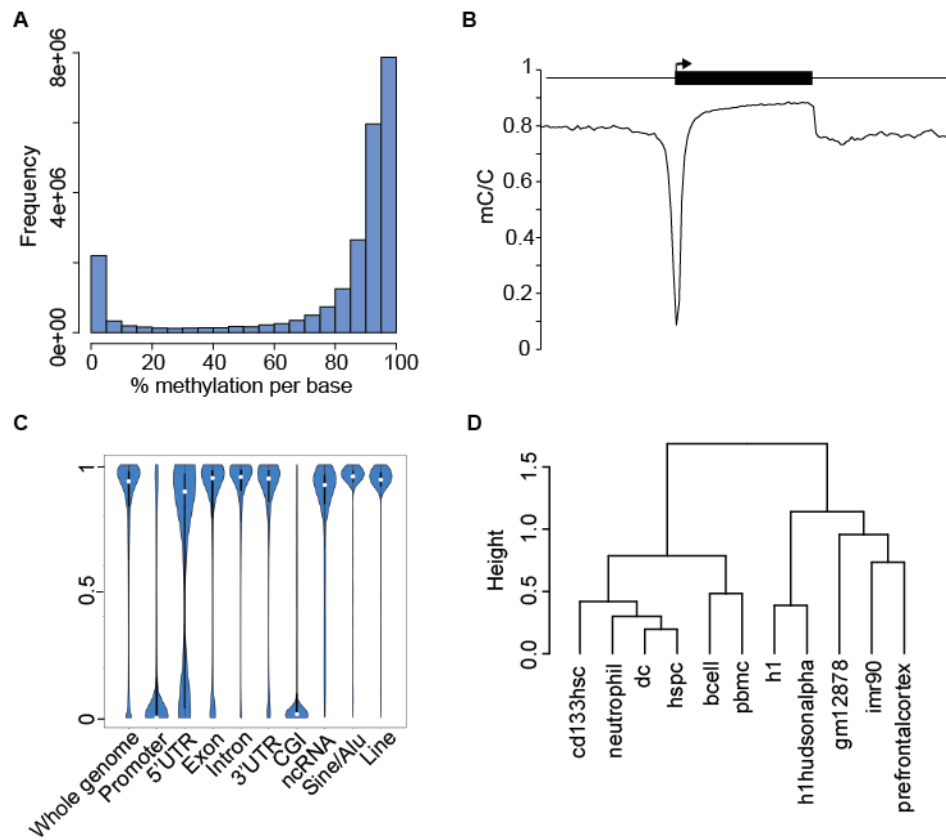


**Figure S1. General Workflow of Study**

	TB81	TB82	TB83	TB87	TB89	TB91	NI81	NI82	NI83	NI87	NI89	NI91
TB81		0.87	0.87	0.88	0.84	0.85	0.91	0.87	0.87	0.86	0.86	0.87
TB82	0.87		0.87	0.88	0.84	0.85	0.87	0.91	0.86	0.86	0.86	0.87
TB83	0.87	0.87		0.88	0.84	0.85	0.87	0.87	0.90	0.86	0.86	0.87
TB87	0.88	0.88	0.88		0.85	0.86	0.88	0.88	0.87	0.91	0.87	0.88
TB89	0.84	0.84	0.84	0.85		0.82	0.84	0.84	0.83	0.83	0.86	0.84
TB91	0.85	0.85	0.85	0.86	0.82		0.86	0.85	0.85	0.85	0.84	0.88
NI81	0.91	0.87	0.87	0.88	0.84	0.86		0.87	0.87	0.87	0.87	0.88
NI82	0.87	0.91	0.87	0.88	0.84	0.85	0.87		0.86	0.86	0.86	0.87
NI83	0.87	0.86	0.90	0.87	0.83	0.85	0.87	0.86		0.86	0.86	0.87
NI87	0.86	0.86	0.86	0.91	0.83	0.85	0.87	0.86	0.86		0.85	0.87
NI89	0.86	0.86	0.86	0.87	0.86	0.84	0.87	0.86	0.86	0.85		0.86
NI91	0.87	0.87	0.87	0.88	0.84	0.88	0.88	0.87	0.87	0.87	0.86	

**Table S1. Correlation of Methylation Levels Between DC Samples**

Correlation coefficients were obtained using only CpG sites covered at least in each sample and have a minimum coverage of 5.



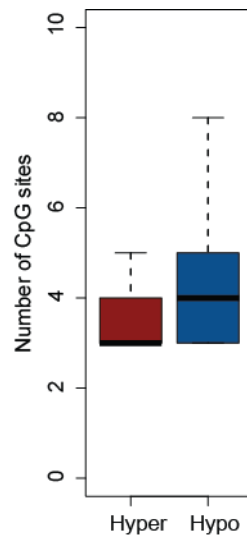
**Figure S2. Global Trends of DNA Methylation in DCs**

(A) Distribution of genome-wide CpG methylation.

(B) Relative CpG methylation level with respect to gene-associated regions.

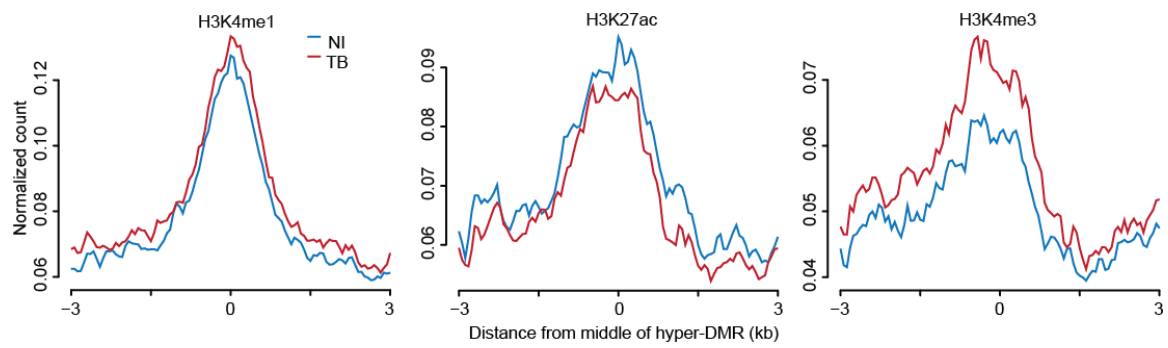
(C) Violin plots indicate the methylation ratios in regions with different genomic features.

(D) Hierarchical clustering using Ward's method based on CpG methylation levels across different cell types. Only CpG sites covered at least in each sample and have a minimum coverage of 5 were used.



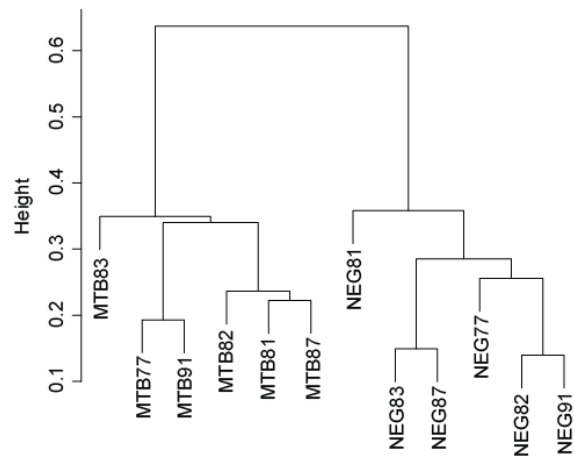
**Figure S3. Genomic Features of MTB-DMRs**

Boxplots comparing the distribution of the number of CpG sites within hyper- and hypomethylated regions.



**Figure S4. Chromatin Dynamics at Hypermethylated Regions**

Composite plots of patterns of histone marks before and after MTB infection, 3 kb around midpoints of hyper-DMRs.



**Figure S5. Transcriptional Profiles of DC Samples**

Hierarchical clustering of DC samples based on whole-genome gene expression data.

## 4 Conclusion

Recent advances in epigenomic profiling technologies have expanded our appreciation of the importance of dynamic epigenetic mechanisms in regulating transcriptional responses to immune challenges. While such epigenetic changes have been well-characterized in patterns of different histone marks, it is still not clear to which extent active changes in DNA methylation participate in immune responses to immune stimuli. Here, we investigated dynamic epigenetic and transcriptional profiles of primary dendritic cell (DC) in response to *Mycobacterium tuberculosis* (MTB) infection. Our results demonstrate that bacterial infection rapidly reconfigures the DNA methylation landscape of DCs. Dynamic alterations in DNA methylation primarily occur at putative enhancers marked by H3K4me1. Hypomethylated regions substantially gain both H3K4me1 and H3K27ac indicative of enhancers becoming active after infection. Although we did not observe such dynamic interplay in hypermethylated regions, this still suggests that there is a close mechanistic link between DNA methylation and histone modifications which is essential in modulating enhancer activity and thus gene expression. Consistent with this, we found that dynamic DNA methylation is associated with genes that are differentially expressed after MTB infection. In concert, these findings provide evidence that changes in DNA methylation at distal regulatory regions are potentially implicated in regulating immune responses to bacterial infection





## Bibliography

- Alvarez, D., E. H. Vollmann and U. H. von Andrian, *Mechanisms and consequences of dendritic cell migration*. Immunity, 2008. 29(3): 325-342.
- Anders, S., D. J. McCarthy, Y. Chen, M. Okoniewski, G. K. Smyth, W. Huber and M. D. Robinson, *Count-based differential expression analysis of RNA sequencing data using R and Bioconductor*. Nat Protoc, 2013. 8(9): 1765-1786.
- Aran, D., G. Toperoff, M. Rosenberg and A. Hellman, *Replication timing-related and gene body-specific methylation of active human genes*. Hum Mol Genet, 2011. 20(4): 670-680.
- Ball, M. P., J. B. Li, Y. Gao, J. H. Lee, E. M. LeProust, I. H. Park, B. Xie, G. Q. Daley and G. M. Church, *Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells*. Nat Biotechnol, 2009. 27(4): 361-368.
- Banchereau, J. and R. M. Steinman, *Dendritic cells and the control of immunity*. Nature, 1998. 392(6673): 245-252.
- Barish, G. D., R. T. Yu, M. Karunasiri, C. B. Ocampo, J. Dixon, C. Benner, A. L. Dent, R. K. Tangirala and R. M. Evans, *Bcl-6 and NF-kappaB cistromes mediate opposing regulation of the innate immune response*. Genes Dev, 2010. 24(24): 2760-2765.
- Barreiro, L. B., L. Tailleux, A. A. Pai, B. Gicquel, J. C. Marioni and Y. Gilad, *Deciphering the genetic architecture of variation in the immune response to Mycobacterium tuberculosis infection*. Proc Natl Acad Sci U S A, 2012. 109(4): 1204-1209.
- Berger, S. L., *The complex language of chromatin regulation during transcription*. Nature, 2007. 447(7143): 407-412.
- Bestor, T., A. Laudano, R. Mattaliano and V. Ingram, *Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases*. J Mol Biol, 1988. 203(4): 971-983.
- Bestor, T. H., *The DNA methyltransferases of mammals*. Hum Mol Genet, 2000. 9(16): 2395-2402.
- Bhutani, N., D. M. Burns and H. M. Blau, *DNA demethylation dynamics*. Cell, 2011. 146(6): 866-872.
- Biddie, S. C., S. John, P. J. Sabo, R. E. Thurman, T. A. Johnson, R. L. Schiltz, T. B. Miranda, M. H. Sung, S. Trump, S. L. Lightman, C. Vinson, J. A. Stamatoyannopoulos and G. L. Hager, *Transcription factor API potentiates chromatin accessibility and glucocorticoid receptor binding*. Mol Cell, 2011. 43(1): 145-155.
- Bird, A., M. Taggart, M. Frommer, O. J. Miller and D. Macleod, *A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA*. Cell, 1985. 40(1): 91-99.
- Bock, C., *Analysing and interpreting DNA methylation data*. Nat Rev Genet, 2012. 13(10): 705-719.
- Bogdanovic, O. and G. J. Veenstra, *DNA methylation and methyl-CpG binding proteins: developmental requirements and function*. Chromosoma, 2009. 118(5): 549-565.
- Broske, A. M., L. Vockentanz, S. Kharazi, M. R. Huska, E. Mancini, M. Scheller, C. Kuhl, A. Enns, M. Prinz, R. Jaenisch, C. Nerlov, A. Leutz, M. A. Andrade-Navarro, S. E. Jacobsen and

F. Rosenbauer, *DNA methylation protects hematopoietic stem cell multipotency from myeloerythroid restriction*. Nat Genet, 2009. 41(11): 1207-1215.

Buenrostro, J. D., P. G. Giresi, L. C. Zaba, H. Y. Chang and W. J. Greenleaf, *Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position*. Nat Methods, 2013. 10(12): 1213-1218.

Cedar, H. and Y. Bergman, *Linking DNA methylation and histone modification: patterns and paradigms*. Nat Rev Genet, 2009. 10(5): 295-304.

Choi, J. K., *Contrasting chromatin organization of CpG islands and exons in the human genome*. Genome Biol, 2010. 11(7): R70.

Clark, S. J., J. Harrison, C. L. Paul and M. Frommer, *High sensitivity mapping of methylated cytosines*. Nucleic Acids Res, 1994. 22(15): 2990-2997.

Clark, S. J., A. Statham, C. Stirzaker, P. L. Molloy and M. Frommer, *DNA methylation: bisulphite modification and analysis*. Nat Protoc, 2006. 1(5): 2353-2364.

Consortium, E. P., B. E. Bernstein, E. Birney, I. Dunham, E. D. Green, C. Gunter and M. Snyder, *An integrated encyclopedia of DNA elements in the human genome*. Nature, 2012. 489(7414): 57-74.

Cooper, A. M., *Cell-mediated immune responses in tuberculosis*. Annu Rev Immunol, 2009. 27: 393-422.

de Saint-Vis, B., I. Fugier-Vivier, C. Massacrier, C. Gaillard, B. Vanbervliet, S. Ait-Yahia, J. Banchereau, Y. J. Liu, S. Lebecque and C. Caux, *The cytokine profile expressed by human dendritic cells is dependent on cell subtype and mode of activation*. J Immunol, 1998. 160(4): 1666-1676.

Deaton, A. M. and A. Bird, *CpG islands and the regulation of transcription*. Genes Dev, 2011. 25(10): 1010-1022.

Eden, E., R. Navon, I. Steinfeld, D. Lipson and Z. Yakhini, *GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists*. BMC Bioinformatics, 2009. 10: 48.

Ernst, J. and M. Kellis, *ChromHMM: automating chromatin-state discovery and characterization*. Nat Methods, 2012. 9(3): 215-216.

Filion, G. J., S. Zhenilo, S. Salozhin, D. Yamada, E. Prokhortchouk and P. A. Defossez, *A family of human zinc finger proteins that bind methylated DNA and repress transcription*. Mol Cell Biol, 2006. 26(1): 169-181.

Fischle, W., *Talk is cheap--cross-talk in establishment, maintenance, and readout of chromatin modifications*. Genes Dev, 2008. 22(24): 3375-3382.

Foster, S. L., D. C. Hargreaves and R. Medzhitov, *Gene-specific control of inflammation by TLR-induced chromatin modifications*. Nature, 2007. 447(7147): 972-978.

Fouse, S. D., Y. Shen, M. Pellegrini, S. Cole, A. Meissner, L. Van Neste, R. Jaenisch and G. Fan, *Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation*. Cell Stem Cell, 2008. 2(2): 160-169.

Fromme, J. C. and G. L. Verdine, *Base excision repair*. Adv Protein Chem, 2004. 69: 1-41.

Garber, M., N. Yosef, A. Goren, R. Raychowdhury, A. Thielke, M. Guttman, J. Robinson, B. Minie, N. Chevrier, Z. Itzhaki, R. Blecher-Gonen, C. Bornstein, D. Amann-Zalcenstein, A. Weiner, D. Friedrich, J. Meldrim, O. Ram, C. Cheng, A. Gnirke, S. Fisher, N. Friedman, B. Wong, B. E. Bernstein, C. Nusbaum, N. Hacohen, A. Regev and I. Amit, *A high-throughput*

*chromatin immunoprecipitation approach reveals principles of dynamic gene regulation in mammals*. Mol Cell, 2012. 47(5): 810-822.

Gardiner-Garden, M. and M. Frommer, *CpG islands in vertebrate genomes*. J Mol Biol, 1987. 196(2): 261-282.

Ghisletti, S., I. Barozzi, F. Mietton, S. Polletti, F. De Santa, E. Venturini, L. Gregory, L. Lonie, A. Chew, C. L. Wei, J. Ragoussis and G. Natoli, *Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages*. Immunity, 2010. 32(3): 317-328.

Gifford, C. A., M. J. Ziller, H. Gu, C. Trapnell, J. Donaghey, A. Tsankov, A. K. Shalek, D. R. Kelley, A. A. Shishkin, R. Issner, X. Zhang, M. Coyne, J. L. Fostel, L. Holmes, J. Meldrim, M. Guttman, C. Epstein, H. Park, O. Kohlbacher, J. Rinn, A. Gnirke, E. S. Lander, B. E. Bernstein and A. Meissner, *Transcriptional and epigenetic dynamics during specification of human embryonic stem cells*. Cell, 2013. 153(5): 1149-1163.

Goll, M. G. and T. H. Bestor, *Eukaryotic cytosine methyltransferases*. Annu Rev Biochem, 2005. 74: 481-514.

Goyal, R., R. Reinhardt and A. Jeltsch, *Accuracy of DNA methylation pattern preservation by the Dnmt1 methyltransferase*. Nucleic Acids Res, 2006. 34(4): 1182-1188.

Guo, J. U., D. K. Ma, H. Mo, M. P. Ball, M. H. Jang, M. A. Bonaguidi, J. A. Balazer, H. L. Eaves, B. Xie, E. Ford, K. Zhang, G. L. Ming, Y. Gao and H. Song, *Neuronal activity modifies the DNA methylation landscape in the adult brain*. Nat Neurosci, 2011. 14(10): 1345-1351.

Guo, J. U., Y. Su, J. H. Shin, J. Shin, H. Li, B. Xie, C. Zhong, S. Hu, T. Le, G. Fan, H. Zhu, Q. Chang, Y. Gao, G. L. Ming and H. Song, *Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain*. Nat Neurosci, 2014. 17(2): 215-222.

Hackstein, H. and A. W. Thomson, *Dendritic cells: emerging pharmacological targets of immunosuppressive drugs*. Nat Rev Immunol, 2004. 4(1): 24-34.

Hansen, K. D., B. Langmead and R. A. Irizarry, *BSmooth: from whole genome bisulfite sequencing reads to differentially methylated regions*. Genome Biol, 2012. 13(10): R83.

Hark, A. T., C. J. Schoenherr, D. J. Katz, R. S. Ingram, J. M. LeVorse and S. M. Tilghman, *CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus*. Nature, 2000. 405(6785): 486-489.

Harris, R. A., T. Wang, C. Coarfa, R. P. Nagarajan, C. Hong, S. L. Downey, B. E. Johnson, S. D. Fouse, A. Delaney, Y. Zhao, A. Olshen, T. Ballinger, X. Zhou, K. J. Forsberg, J. Gu, L. Echipare, H. O'Geen, R. Lister, M. Pelizzola, Y. Xi, C. B. Epstein, B. E. Bernstein, R. D. Hawkins, B. Ren, W. Y. Chung, H. Gu, C. Bock, A. Gnirke, M. Q. Zhang, D. Haussler, J. R. Ecker, W. Li, P. J. Farnham, R. A. Waterland, A. Meissner, M. A. Marra, M. Hirst, A. Milosavljevic and J. F. Costello, *Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications*. Nat Biotechnol, 2010. 28(10): 1097-1105.

Hashimoto, H., J. R. Horton, X. Zhang, M. Bostick, S. E. Jacobsen and X. Cheng, *The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix*. Nature, 2008. 455(7214): 826-829.

Hashimoto, H., J. R. Horton, X. Zhang and X. Cheng, *UHRF1, a modular multi-domain protein, regulates replication-coupled crosstalk between DNA methylation and histone modifications*. Epigenetics, 2009. 4(1): 8-14.

Hayden, M. S. and S. Ghosh, *Shared principles in NF-kappaB signaling*. Cell, 2008. 132(3): 344-362.

He, Y. F., B. Z. Li, Z. Li, P. Liu, Y. Wang, Q. Tang, J. Ding, Y. Jia, Z. Chen, L. Li, Y. Sun, X. Li, Q. Dai, C. X. Song, K. Zhang, C. He and G. L. Xu, *Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA*. Science, 2011. 333(6047): 1303-1307.

Heintzman, N. D., G. C. Hon, R. D. Hawkins, P. Kheradpour, A. Stark, L. F. Harp, Z. Ye, L. K. Lee, R. K. Stuart, C. W. Ching, K. A. Ching, J. E. Antosiewicz-Bourget, H. Liu, X. Zhang, R. D. Green, V. V. Lobanenko, R. Stewart, J. A. Thomson, G. E. Crawford, M. Kellis and B. Ren, *Histone modifications at human enhancers reflect global cell-type-specific gene expression*. Nature, 2009. 459(7243): 108-112.

Hellman, A. and A. Chess, *Gene body-specific methylation on the active X chromosome*. Science, 2007. 315(5815): 1141-1143.

Hodges, E., A. Molaro, C. O. Dos Santos, P. Thekkat, Q. Song, P. J. Uren, J. Park, J. Butler, S. Rafii, W. R. McCombie, A. D. Smith and G. J. Hannon, *Directional DNA methylation changes and complex intermediate states accompany lineage specificity in the adult hematopoietic compartment*. Mol Cell, 2011. 44(1): 17-28.

Hogart, A., J. Lichtenberg, S. S. Ajay, S. Anderson, N. I. H. I. S. Center, E. H. Margulies and D. M. Bodine, *Genome-wide DNA methylation profiles in hematopoietic stem and progenitor cells reveal overrepresentation of ETS transcription factor binding sites*. Genome Res, 2012. 22(8): 1407-1418.

Hon, G. C., N. Rajagopal, Y. Shen, D. F. McCleary, F. Yue, M. D. Dang and B. Ren, *Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues*. Nat Genet, 2013. 45(10): 1198-1206.

Illingworth, R. S., U. Gruenewald-Schneider, S. Webb, A. R. Kerr, K. D. James, D. J. Turner, C. Smith, D. J. Harrison, R. Andrews and A. P. Bird, *Orphan CpG islands identify numerous conserved promoters in the mammalian genome*. PLoS Genet, 2010. 6(9): e1001134.

Inoue, A., L. Shen, Q. Dai, C. He and Y. Zhang, *Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development*. Cell Res, 2011. 21(12): 1670-1676.

Inoue, A. and Y. Zhang, *Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos*. Science, 2011. 334(6053): 194.

Ito, S., L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He and Y. Zhang, *Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine*. Science, 2011. 333(6047): 1300-1303.

Jaenisch, R. and A. Bird, *Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals*. Nat Genet, 2003. 33 Suppl: 245-254.

Janeway, C. A., Jr. and R. Medzhitov, *Innate immune recognition*. Annu Rev Immunol, 2002. 20: 197-216.

Jia, D., R. Z. Jurkowska, X. Zhang, A. Jeltsch and X. Cheng, *Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation*. Nature, 2007. 449(7159): 248-251.

John, S., P. J. Sabo, T. K. Canfield, K. Lee, S. Vong, M. Weaver, H. Wang, J. Vierstra, A. P. Reynolds, R. E. Thurman and J. A. Stamatoyannopoulos, *Genome-scale mapping of DNase I hypersensitivity*. Curr Protoc Mol Biol, 2013. Chapter 27: Unit 21 27.

Jones, P. A. and S. B. Baylin, *The fundamental role of epigenetic events in cancer*. Nat Rev Genet, 2002. 3(6): 415-428.

Jones, P. A., S. M. Taylor and V. L. Wilson, *Inhibition of DNA methylation by 5-azacytidine*. Recent Results Cancer Res, 1983. 84: 202-211.

Kaikkonen, M. U., N. J. Spann, S. Heinz, C. E. Romanoski, K. A. Allison, J. D. Stender, H. B. Chun, D. F. Tough, R. K. Prinjha, C. Benner and C. K. Glass, *Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription*. Mol Cell, 2013. 51(3): 310-325.

Kaneda, M., M. Okano, K. Hata, T. Sado, N. Tsujimoto, E. Li and H. Sasaki, *Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting*. Nature, 2004. 429(6994): 900-903.

Kapsenberg, M. L., *Dendritic-cell control of pathogen-driven T-cell polarization*. Nat Rev Immunol, 2003. 3(12): 984-993.

Kim, D., G. Pertea, C. Trapnell, H. Pimentel, R. Kelley and S. L. Salzberg, *TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions*. Genome Biol, 2013. 14(4): R36.

Kim, T. K., M. Hemberg, J. M. Gray, A. M. Costa, D. M. Bear, J. Wu, D. A. Harmin, M. Laptewicz, K. Barbara-Haley, S. Kuersten, E. Markenscoff-Papadimitriou, D. Kuhl, H. Bito, P. F. Worley, G. Kreiman and M. E. Greenberg, *Widespread transcription at neuronal activity-regulated enhancers*. Nature, 2010. 465(7295): 182-187.

Klug, M., S. Heinz, C. Gebhard, L. Schwarzfischer, S. W. Krause, R. Andreessen and M. Rehli, *Active DNA demethylation in human postmitotic cells correlates with activating histone modifications, but not transcription levels*. Genome Biol, 2010. 11(6): R63.

Kohli, R. M. and Y. Zhang, *TET enzymes, TDG and the dynamics of DNA demethylation*. Nature, 2013. 502(7472): 472-479.

Kremer, L., A. Baulard, J. Estaquier, O. Poulain-Godefroy and C. Locht, *Green fluorescent protein as a new expression marker in mycobacteria*. Mol Microbiol, 1995. 17(5): 913-922.

Krueger, F. and S. R. Andrews, *Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications*. Bioinformatics, 2011. 27(11): 1571-1572.

Laird, P. W., *Principles and challenges of genomewide DNA methylation analysis*. Nat Rev Genet, 2010. 11(3): 191-203.

Lander, E. S., L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. LeHoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J. P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J. C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R. H. Waterston, R. K. Wilson, L. W. Hillier, J. D. McPherson, M. A. Marra, E. R. Mardis, L. A. Fulton, A. T. Chinwalla, K. H. Pepin, W. R. Gish, S. L. Chisoe, M. C. Wendl, K. D. Delehaunty, T. L. Miner, A. Delehaunty, J. B. Kramer, L. L. Cook, R. S. Fulton, D. L. Johnson, P. J. Minx, S. W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J. F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R. A. Gibbs, D.

M. Muzny, S. E. Scherer, J. B. Bouck, E. J. Sodergren, K. C. Worley, C. M. Rives, J. H. Gorrell, M. L. Metzker, S. L. Naylor, R. S. Kucherlapati, D. L. Nelson, G. M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, D. R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, H. M. Lee, J. Dubois, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R. W. Davis, N. A. Federspiel, A. P. Abola, M. J. Proctor, R. M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D. R. Cox, M. V. Olson, R. Kaul, C. Raymond, N. Shimizu, K. Kawasaki, S. Minoshima, G. A. Evans, M. Athanasiou, R. Schultz, B. A. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W. R. McCombie, M. de la Bastide, N. Dedhia, H. Blocker, K. Hornischer, G. Nordsiek, R. Agarwala, L. Aravind, J. A. Bailey, A. Bateman, S. Batzoglu, E. Birney, P. Bork, D. G. Brown, C. B. Burge, L. Cerutti, H. C. Chen, D. Church, M. Clamp, R. R. Copley, T. Doerks, S. R. Eddy, E. E. Eichler, T. S. Furey, J. Galagan, J. G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L. S. Johnson, T. A. Jones, S. Kasif, A. Kasprzyk, S. Kennedy, W. J. Kent, P. Kitts, E. V. Koonin, I. Korf, D. Kulp, D. Lancet, T. M. Lowe, A. McLysaght, T. Mikkelsen, J. V. Moran, N. Mulder, V. J. Pollara, C. P. Ponting, G. Schuler, J. Schultz, G. Slater, A. F. Smit, E. Stupka, J. Szustakowski, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y. I. Wolf, K. H. Wolfe, S. P. Yang, R. F. Yeh, F. Collins, M. S. Guyer, J. Peterson, A. Felsenfeld, K. A. Wetterstrand, A. Patrinos, M. J. Morgan, P. de Jong, J. J. Catanese, K. Osoegawa, H. Shizuya, S. Choi, Y. J. Chen and C. International Human Genome Sequencing, *Initial sequencing and analysis of the human genome*. Nature, 2001. 409(6822): 860-921.

Langmead, B. and S. L. Salzberg, *Fast gapped-read alignment with Bowtie 2*. Nat Methods, 2012. 9(4): 357-359.

Laurent, L., E. Wong, G. Li, T. Huynh, A. Tsirigos, C. T. Ong, H. M. Low, K. W. Kin Sung, I. Rigoutsos, J. Loring and C. L. Wei, *Dynamic changes in the human methylome during differentiation*. Genome Res, 2010. 20(3): 320-331.

Leonhardt, H., A. W. Page, H. U. Weier and T. H. Bestor, *A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei*. Cell, 1992. 71(5): 865-873.

Li, E., C. Beard and R. Jaenisch, *Role for DNA methylation in genomic imprinting*. Nature, 1993. 366(6453): 362-365.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin and S. Genome Project Data Processing, *The Sequence Alignment/Map format and SAMtools*. Bioinformatics, 2009. 25(16): 2078-2079.

Lister, R., E. A. Mukamel, J. R. Nery, M. Urich, C. A. Puddifoot, N. D. Johnson, J. Lucero, Y. Huang, A. J. Dwork, M. D. Schultz, M. Yu, J. Tonti-Filippini, H. Heyn, S. Hu, J. C. Wu, A. Rao, M. Esteller, C. He, F. G. Haghghi, T. J. Sejnowski, M. M. Behrens and J. R. Ecker, *Global epigenomic reconfiguration during mammalian brain development*. Science, 2013. 341(6146): 1237905.

Lister, R., M. Pelizzola, R. H. Dowen, R. D. Hawkins, G. Hon, J. Tonti-Filippini, J. R. Nery, L. Lee, Z. Ye, Q. M. Ngo, L. Edsall, J. Antosiewicz-Bourget, R. Stewart, V. Ruotti, A. H. Millar, J. A. Thomson, B. Ren and J. R. Ecker, *Human DNA methylomes at base resolution show widespread epigenomic differences*. Nature, 2009. 462(7271): 315-322.

Martinez-Balbas, M. A., A. J. Bannister, K. Martin, P. Haus-Seuffert, M. Meisterernst and T. Kouzarides, *The acetyltransferase activity of CBP stimulates transcription*. EMBO J, 1998. 17(10): 2886-2893.

Maunakea, A. K., I. Chepelev, K. Cui and K. Zhao, *Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition*. Cell Res, 2013. 23(11): 1256-1269.

Maunakea, A. K., R. P. Nagarajan, M. Bilenky, T. J. Ballinger, C. D'Souza, S. D. Fouse, B. E. Johnson, C. Hong, C. Nielsen, Y. Zhao, G. Turecki, A. Delaney, R. Varhol, N. Thiessen, K. Shchors, V. M. Heine, D. H. Rowitch, X. Xing, C. Fiore, M. Schillebeeckx, S. J. Jones, D. Haussler, M. A. Marra, M. Hirst, T. Wang and J. F. Costello, *Conserved role of intragenic DNA methylation in regulating alternative promoters*. Nature, 2010. 466(7303): 253-257.

McGhee, J. D. and G. D. Ginder, *Specific DNA methylation sites in the vicinity of the chicken beta-globin genes*. Nature, 1979. 280(5721): 419-420.

Medzhitov, R., *Toll-like receptors and innate immunity*. Nat Rev Immunol, 2001. 1(2): 135-145.

Medzhitov, R., *TLR-mediated innate immune recognition*. Semin Immunol, 2007. 19(1): 1-2.

Medzhitov, R. and T. Horng, *Transcriptional control of the inflammatory response*. Nat Rev Immunol, 2009. 9(10): 692-703.

Meehan, R. R., J. D. Lewis, S. McKay, E. L. Kleiner and A. P. Bird, *Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs*. Cell, 1989. 58(3): 499-507.

Meissner, A., T. S. Mikkelsen, H. Gu, M. Wernig, J. Hanna, A. Sivachenko, X. Zhang, B. E. Bernstein, C. Nusbaum, D. B. Jaffe, A. Gnirke, R. Jaenisch and E. S. Lander, *Genome-scale DNA methylation maps of pluripotent and differentiated cells*. Nature, 2008. 454(7205): 766-770.

Nan, X., R. R. Meehan and A. Bird, *Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2*. Nucleic Acids Res, 1993. 21(21): 4886-4892.

Neph, S., J. Vierstra, A. B. Stergachis, A. P. Reynolds, E. Haugen, B. Vernot, R. E. Thurman, S. John, R. Sandstrom, A. K. Johnson, M. T. Maurano, R. Humbert, E. Rynes, H. Wang, S. Vong, K. Lee, D. Bates, M. Diegel, V. Roach, D. Dunn, J. Neri, A. Schafer, R. S. Hansen, T. Kutayavin, E. Giste, M. Weaver, T. Canfield, P. Sabo, M. Zhang, G. Balasundaram, R. Byron, M. J. MacCoss, J. M. Akey, M. A. Bender, M. Groudine, R. Kaul and J. A. Stamatoyannopoulos, *An expansive human regulatory lexicon encoded in transcription factor footprints*. Nature, 2012. 489(7414): 83-90.

Ogryzko, V. V., R. L. Schiltz, V. Russanova, B. H. Howard and Y. Nakatani, *The transcriptional coactivators p300 and CBP are histone acetyltransferases*. Cell, 1996. 87(5): 953-959.

Okano, M., D. W. Bell, D. A. Haber and E. Li, *DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development*. Cell, 1999. 99(3): 247-257.

Ooi, S. K., C. Qiu, E. Bernstein, K. Li, D. Jia, Z. Yang, H. Erdjument-Bromage, P. Tempst, S. P. Lin, C. D. Allis, X. Cheng and T. H. Bestor, *DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA*. Nature, 2007. 448(7154): 714-717.

Ostuni, R., V. Piccolo, I. Barozzi, S. Polletti, A. Termanini, S. Bonifacio, A. Curina, E. Prosperini, S. Ghisletti and G. Natoli, *Latent enhancers activated by stimulation in differentiated cells*. Cell, 2013. 152(1-2): 157-171.

Pique-Regi, R., J. F. Degner, A. A. Pai, D. J. Gaffney, Y. Gilad and J. K. Pritchard, *Accurate inference of transcription factor binding from DNA sequence and chromatin accessibility data*. Genome Res, 2011. 21(3): 447-455.

Prokhortchouk, A., B. Hendrich, H. Jorgensen, A. Ruzov, M. Wilm, G. Georgiev, A. Bird and E. Prokhortchouk, *The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor*. Genes Dev, 2001. 15(13): 1613-1618.

Quinlan, A. R. and I. M. Hall, *BEDTools: a flexible suite of utilities for comparing genomic features*. Bioinformatics, 2010. 26(6): 841-842.

Quintin, J., S. C. Cheng, J. W. van der Meer and M. G. Netea, *Innate immune memory: towards a better understanding of host defense mechanisms*. Curr Opin Immunol, 2014. 29C: 1-7.

Ramirez-Carrozzi, V. R., D. Braas, D. M. Bhatt, C. S. Cheng, C. Hong, K. R. Doty, J. C. Black, A. Hoffmann, M. Carey and S. T. Smale, *A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling*. Cell, 2009. 138(1): 114-128.

Ranz, J. M., C. I. Castillo-Davis, C. D. Meiklejohn and D. L. Hartl, *Sex-dependent gene expression and evolution of the Drosophila transcriptome*. Science, 2003. 300(5626): 1742-1745.

Razin, A., *CpG methylation, chromatin structure and gene silencing-a three-way connection*. EMBO J, 1998. 17(17): 4905-4908.

Reik, W., G. Kelsey and J. Walter, *Dissecting de novo methylation*. Nat Genet, 1999. 23(4): 380-382.

Rinn, J. L. and M. Snyder, *Sexual dimorphism in mammalian gene expression*. Trends Genet, 2005. 21(5): 298-305.

Robertson, K. D., *DNA methylation and human disease*. Nat Rev Genet, 2005. 6(8): 597-610.

Saxonov, S., P. Berg and D. L. Brutlag, *A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters*. Proc Natl Acad Sci U S A, 2006. 103(5): 1412-1417.

Schermelleh, L., A. Haemmer, F. Spada, N. Rosing, D. Meilinger, U. Rothbauer, M. C. Cardoso and H. Leonhardt, *Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation*. Nucleic Acids Res, 2007. 35(13): 4301-4312.

Schermelleh, L., F. Spada, H. P. Easwaran, K. Zolghadr, J. B. Margot, M. C. Cardoso and H. Leonhardt, *Trapped in action: direct visualization of DNA methyltransferase activity in living cells*. Nat Methods, 2005. 2(10): 751-756.

Schlesinger, F., A. D. Smith, T. R. Gingeras, G. J. Hannon and E. Hodges, *De novo DNA demethylation and noncoding transcription define active intergenic regulatory elements*. Genome Res, 2013. 23(10): 1601-1614.

Shen, L., N. Shao, X. Liu and E. Nestler, *ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases*. BMC Genomics, 2014. 15(1): 284.



Shirane, K., H. Toh, H. Kobayashi, F. Miura, H. Chiba, T. Ito, T. Kono and H. Sasaki, *Mouse oocyte methylomes at base resolution reveal genome-wide accumulation of non-CpG methylation and role of DNA methyltransferases*. PLoS Genet, 2013. 9(4): e1003439.

Siedlecki, P. and P. Zielenkiewicz, *Mammalian DNA methyltransferases*. Acta Biochim Pol, 2006. 53(2): 245-256.

Smale, S. T., *Selective transcription in response to an inflammatory stimulus*. Cell, 2010. 140(6): 833-844.

Smale, S. T., *Hierarchies of NF-kappaB target-gene regulation*. Nat Immunol, 2011. 12(8): 689-694.

Smale, S. T., A. Tarakhovsky and G. Natoli, *Chromatin contributions to the regulation of innate immunity*. Annu Rev Immunol, 2014. 32: 489-511.

Sorensen, A. L. and P. Collas, *Immunoprecipitation of methylated DNA*. Methods Mol Biol, 2009. 567: 249-262.

Stadler, M. B., R. Murr, L. Burger, R. Ivanek, F. Lienert, A. Scholer, E. van Nimwegen, C. Wirbelauer, E. J. Oakeley, D. Gaidatzis, V. K. Tiwari and D. Schubeler, *DNA-binding factors shape the mouse methylome at distal regulatory regions*. Nature, 2011. 480(7378): 490-495.

Tahiliani, M., K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind and A. Rao, *Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1*. Science, 2009. 324(5929): 930-935.

Tailleux, L., O. Neyrolles, S. Honore-Bouakline, E. Perret, F. Sanchez, J. P. Abastado, P. H. Lagrange, J. C. Gluckman, M. Rosenzweig and J. L. Herrmann, *Constrained intracellular survival of Mycobacterium tuberculosis in human dendritic cells*. J Immunol, 2003. 170(4): 1939-1948.

Tailleux, L., S. J. Waddell, M. Pelizzola, A. Mortellaro, M. Withers, A. Tanne, P. R. Castagnoli, B. Gicquel, N. G. Stoker, P. D. Butcher, M. Foti and O. Neyrolles, *Probing host pathogen cross-talk by transcriptional profiling of both Mycobacterium tuberculosis and infected human dendritic cells and macrophages*. PLoS One, 2008. 3(1): e1403.

Tan, J. K. and H. C. O'Neill, *Maturation requirements for dendritic cells in T cell stimulation leading to tolerance versus immunity*. J Leukoc Biol, 2005. 78(2): 319-324.

Taniguchi, T., K. Ogasawara, A. Takaoka and N. Tanaka, *IRF family of transcription factors as regulators of host defense*. Annu Rev Immunol, 2001. 19: 623-655.

Tanne, A., B. Ma, F. Boudou, L. Tailleux, H. Botella, E. Badell, F. Levillain, M. E. Taylor, K. Drickamer, J. Nigou, K. M. Dobos, G. Puzo, D. Vestweber, M. K. Wild, M. Marcinko, P. Sobieszczuk, L. Stewart, D. Lebus, B. Gicquel and O. Neyrolles, *A murine DC-SIGN homologue contributes to early host defense against Mycobacterium tuberculosis*. J Exp Med, 2009. 206(10): 2205-2220.

Tazi, J. and A. Bird, *Alternative chromatin structure at CpG islands*. Cell, 1990. 60(6): 909-920.

Thomson, J. P., P. J. Skene, J. Selfridge, T. Clouaire, J. Guy, S. Webb, A. R. Kerr, A. Deaton, R. Andrews, K. D. James, D. J. Turner, R. Illingworth and A. Bird, *CpG islands influence chromatin structure via the CpG-binding protein Cfp1*. Nature, 2010. 464(7291): 1082-1086.

Thurman, R. E., E. Rynes, R. Humbert, J. Vierstra, M. T. Maurano, E. Haugen, N. C. Sheffield, A. B. Stergachis, H. Wang, B. Vernot, K. Garg, S. John, R. Sandstrom, D. Bates, L. Boatman, T. K. Canfield, M. Diegel, D. Dunn, A. K. Ebersol, T. Frum, E. Giste, A. K. Johnson, E. M.

Johnson, T. Kutuyavin, B. Lajoie, B. K. Lee, K. Lee, D. London, D. Lotakis, S. Neph, F. Neri, E. D. Nguyen, H. Qu, A. P. Reynolds, V. Roach, A. Safi, M. E. Sanchez, A. Sanyal, A. Shafer, J. M. Simon, L. Song, S. Vong, M. Weaver, Y. Yan, Z. Zhang, Z. Zhang, B. Lenhard, M. Tewari, M. O. Dorschner, R. S. Hansen, P. A. Navas, G. Stamatoyannopoulos, V. R. Iyer, J. D. Lieb, S. R. Sunyaev, J. M. Akey, P. J. Sabo, R. Kaul, T. S. Furey, J. Dekker, G. E. Crawford and J. A. Stamatoyannopoulos, *The accessible chromatin landscape of the human genome*. Nature, 2012. 489(7414): 75-82.

Trzyna, E., M. Duleba, M. Faryna and M. Majka, *Regulation of transcription in cancer*. Front Biosci (Landmark Ed), 2012. 17: 316-330.

Weber, M., I. Hellmann, M. B. Stadler, L. Ramos, S. Paabo, M. Rebhan and D. Schubeler, *Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome*. Nat Genet, 2007. 39(4): 457-466.

Xie, W., C. L. Barr, A. Kim, F. Yue, A. Y. Lee, J. Eubanks, E. L. Dempster and B. Ren, *Base-resolution analyses of sequence and parent-of-origin dependent DNA methylation in the mouse genome*. Cell, 2012. 148(4): 816-831.

Xu, G. L., T. H. Bestor, D. Bourc'his, C. L. Hsieh, N. Tommerup, M. Bugge, M. Hulten, X. Qu, J. J. Russo and E. Viegas-Pequignot, *Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene*. Nature, 1999. 402(6758): 187-191.

Yoder, J. A., C. P. Walsh and T. H. Bestor, *Cytosine methylation and the ecology of intragenomic parasites*. Trends Genet, 1997. 13(8): 335-340.

Yu, M., G. C. Hon, K. E. Szulwach, C. X. Song, L. Zhang, A. Kim, X. Li, Q. Dai, Y. Shen, B. Park, J. H. Min, P. Jin, B. Ren and C. He, *Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome*. Cell, 2012. 149(6): 1368-1380.

Zhang, Y., T. Liu, C. A. Meyer, J. Eeckhoute, D. S. Johnson, B. E. Bernstein, C. Nusbaum, R. M. Myers, M. Brown, W. Li and X. S. Liu, *Model-based analysis of ChIP-Seq (MACS)*. Genome Biol, 2008. 9(9): R137.

Ziller, M. J., H. Gu, F. Muller, J. Donaghey, L. T. Tsai, O. Kohlbacher, P. L. De Jager, E. D. Rosen, D. A. Bennett, B. E. Bernstein, A. Gnirke and A. Meissner, *Charting a dynamic DNA methylation landscape of the human genome*. Nature, 2013. 500(7463): 477-481.